

MULTIGRID METHODS FOR SADDLE POINT PROBLEMS: DARCY SYSTEMS

SUSANNE C. BRENNER, DUK-SOON OH, AND LI-YENG SUNG

ABSTRACT. We design and analyze multigrid methods for the saddle point problems resulting from Raviart-Thomas-Nédélec mixed finite element methods (of order at least 1) for the Darcy system in porous media flow. Uniform convergence of the W -cycle algorithm in a nonstandard energy norm is established. Extensions to general second order elliptic problems are also addressed.

1. INTRODUCTION

Multigrid methods for saddle point problems arising from mixed finite element methods for Stokes and Lamé systems were investigated in the recent paper [17], where uniform convergence for the W -cycle algorithm in the energy norm was established for arbitrary polyhedral domains. In this paper we will extend the results in [17] to the Darcy system in porous media flow, and to general second order elliptic problems. We will follow the standard notation for differential operators and function spaces that can be found, for example, in [21, 18, 10].

Let Ω be a polyhedral domain in \mathbb{R}^d ($d = 2, 3$) occupied by a porous media. The velocity \mathbf{u} and pressure p of a flow in Ω that obeys Darcy's law are determined by the system of equations

$$(1.1) \quad \mathbf{u} = -\mathbf{A}\nabla p \quad \text{in } \Omega,$$

$$(1.2) \quad \nabla \cdot \mathbf{u} = f \quad \text{in } \Omega,$$

together with the boundary condition

$$(1.3) \quad p = g \quad \text{on } \partial\Omega.$$

Here f is a source, g is the pressure on $\partial\Omega$, and \mathbf{A} , a (sufficiently) smooth $d \times d$ symmetric positive definite (SPD) matrix function on $\bar{\Omega}$, is the permeability tensor divided by the viscosity.

1991 *Mathematics Subject Classification.* 65N55, 65N30, 65N15.

Key words and phrases. multigrid, saddle point problem, Darcy, second order elliptic problems, Raviart-Thomas-Nédélec mixed finite element methods.

The work of the first and third authors was supported in part by the National Science Foundation under Grant No. DMS-13-19172.

For the design and analysis of multigrid methods, it suffices to consider the case where $g = 0$. A standard weak formulation [10] of (1.1)–(1.3) is then to find $(\mathbf{u}, p) \in H(\operatorname{div}; \Omega) \times L_2(\Omega)$ such that

$$(1.4) \quad a(\mathbf{u}, \mathbf{v}) + b(\mathbf{v}, p) = 0 \quad \forall \mathbf{v} \in H(\operatorname{div}; \Omega),$$

$$(1.5) \quad b(\mathbf{u}, q) = F(q) \quad \forall q \in L_2(\Omega),$$

where

$$a(\mathbf{w}, \mathbf{v}) = \int_{\Omega} (\mathbf{A}^{-1} \mathbf{w}) \cdot \mathbf{v} \, dx, \quad b(\mathbf{v}, q) = - \int_{\Omega} (\nabla \cdot \mathbf{v}) q \, dx \quad \text{and} \quad F(q) = - \int_{\Omega} f q \, dx.$$

Note that (1.4)–(1.5) can be written concisely as

$$(1.6) \quad \mathcal{B}((\mathbf{u}, p), (\mathbf{v}, q)) = F(q) \quad \forall (\mathbf{v}, q) \in H(\operatorname{div}; \Omega) \times L_2(\Omega),$$

where

$$(1.7) \quad \mathcal{B}((\mathbf{w}, r), (\mathbf{v}, q)) = a(\mathbf{w}, \mathbf{v}) + b(\mathbf{v}, r) + b(\mathbf{w}, q).$$

Let \mathcal{T}_h be a simplicial triangulation of Ω . The Raviart-Thomas-Nédélec finite element method [38, 37] for (1.4)–(1.5) is to find $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ such that

$$(1.8) \quad \mathcal{B}((\mathbf{u}_h, p_h), (\mathbf{v}, q)) = F(q) \quad \forall \mathbf{v} \in V_h \times Q_h,$$

where $V_h \subset H(\operatorname{div}; \Omega)$ is the Raviart-Thomas-Nédélec vector finite element space of order $\ell \geq 1$ associated with \mathcal{T}_h and $Q_h \subset L_2(\Omega)$ is the space of discontinuous piecewise P_ℓ functions.

We will consider, as in [17], all-at-once multigrid methods that compute \mathbf{u} and p simultaneously. There is, however, a fundamental difference between the saddle point problems for the Stokes and Lamé systems considered in [17] and the saddle point problem (1.8).

For the saddle point problems in [17], the vector variable belongs to $[H^1(\Omega)]^d$ and the scalar variable belongs to $L_2(\Omega)$, which are the correct spaces for the duality argument that appears in the proof of the approximation properties of the multigrid algorithms. For the saddle point problem (1.8), the vector variable belongs to $H(\operatorname{div}; \Omega)$ and the scalar variable belongs to $L_2(\Omega)$, which are *not* the correct spaces for the duality argument that is based on elliptic regularity (cf. (1.11) below).

This difficulty regarding the saddle point problem defined by (1.8) can be remedied by treating it as a *nonconforming* method for the following alternative weak formulation of the Darcy system: Find $(\mathbf{u}, p) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ such that

$$(1.9) \quad a(\mathbf{u}, \mathbf{v}) + b'(\mathbf{v}, p) = 0 \quad \forall \mathbf{v} \in [L_2(\Omega)]^d,$$

$$(1.10) \quad b'(\mathbf{u}, q) = F(q) \quad \forall q \in H_0^1(\Omega),$$

where

$$b'(\mathbf{v}, q) = \int_{\Omega} \mathbf{v} \cdot \nabla q \, dx.$$

The weak formulation (1.9)–(1.10) is well-defined for $F \in H^{-s}(\Omega)$ for $0 \leq s \leq 1$, and we have the following elliptic regularity estimate (cf. [30, 23, 36]):

$$(1.11) \quad \|\mathbf{u}\|_{H^\alpha(\Omega)} + \|p\|_{H^{1+\alpha}(\Omega)} \leq C_{\Omega, \mathbf{A}} \|F\|_{H^{-1+\alpha}(\Omega)},$$

where $\alpha \in (\frac{1}{2}, 1]$ is determined by Ω and \mathbf{A} , and $\alpha = 1$ if Ω is convex.

Remark 1.1. Note that (1.8) is well-defined for $F \in H^{-s}(\Omega)$ as long as $s < \frac{1}{2}$ since Q_h is a subspace of $H^s(\Omega)$ for any $s < \frac{1}{2}$. Moreover (1.6) remains valid for the solution (\mathbf{u}, p) of (1.9)–(1.10) and for any $(\mathbf{v}, q) \in V_h \times Q_h$, provided we use the following interpretation of $b(\cdot, \cdot)$:

$$b(\mathbf{v}, q) = -\langle \nabla \cdot \mathbf{v}, q \rangle_{H^{-s}(\Omega) \times H^s(\Omega)}$$

where $\langle \cdot, \cdot \rangle_{H^{-s}(\Omega) \times H^s(\Omega)}$ is the canonical bilinear form on $H^{-s}(\Omega) \times H^s(\Omega)$.

The weak formulation defined by (1.9)–(1.10) provides the correct setting for a duality argument based on (1.11), which allows us to establish uniform convergence for W -cycle algorithms for (1.8) in a nonconforming energy norm related to (1.9)–(1.10). This is also the reason that we require the order of the Raviart-Thomas-Nédélec finite element method to be at least 1, since piecewise constant functions provide poor approximations of functions in $H_0^1(\Omega)$ (cf. Remark 2.1).

We note that multigrid algorithms for the lowest order Raviart-Thomas-Nédélec finite element method can be developed through its connection to the Crouzeix-Raviart nonconforming P_1 finite element method [22, 3, 14]. There are also other multilevel iterative solvers for the Darcy system. We refer the readers to [40, 26, 1, 39, 9, 44, 35] for a discussion of such methods.

The rest of the paper is organized as follows. We present the nonstandard error analysis for the Raviart-Thomas-Nédélec finite element method in Section 2. The results in this section are important for the convergence analysis of the multigrid methods and also shed new light on these finite element methods. We introduce the multigrid algorithms in Section 3 and mesh-dependent norms in Section 4, which are important tools for the convergence analysis carried out in Section 5. In Section 6 we extend the results to mixed finite element methods for generalized Darcy systems arising from general second order elliptic problems. Numerical results are presented in Section 7, followed by some concluding remarks in Section 8.

Throughout the paper we will use C (with or without subscripts) to denote a generic positive constant that depends only on the domain Ω , the order ℓ of the finite element spaces and the shape regularity of the triangulations, but not the mesh sizes. To avoid the proliferation of constants, we also use the notation $A \lesssim B$ (or $A \gtrsim B$) to represent $A \leq (\text{generic constant}) \times B$. The notation $A \approx B$ is equivalent to $A \lesssim B$ and $A \gtrsim B$.

2. A NONSTANDARD ERROR ANALYSIS FOR RAVIART-THOMAS-NÉDÉLEC FINITE ELEMENT METHODS

In this section we will carry out the error analysis of (1.8) as a nonconforming finite element method for (1.9)–(1.10). The analysis is based on mesh-dependent norms and the saddle point theory of Babuška [6] and Brezzi [20]. Similar ideas have been applied to the analysis of mixed finite element methods for the biharmonic problem [5].

2.1. Mesh-Dependent Norms for the Finite Element Spaces. The norm $\|\cdot\|_{L_2(\Omega;\mathcal{T}_h)}$ on $[H^\alpha(\Omega)]^d + V_h$ is defined by

$$(2.1) \quad \|\mathbf{v}\|_{L_2(\Omega;\mathcal{T}_h)}^2 = \sum_{T \in \mathcal{T}_h} \|\mathbf{v}\|_{L_2(T)}^2 + \sum_{\sigma \in \mathfrak{S}_h} h_\sigma \|\mathbf{v} \cdot \mathbf{n}_\sigma\|_{L_2(\sigma)}^2,$$

where \mathfrak{S}_h is the set of the sides (faces for $d = 3$ and edges for $d = 2$) of the elements in \mathcal{T}_h , h_σ is the diameter of the side σ , and \mathbf{n}_σ is a unit normal of σ .

Note that

$$(2.2) \quad \|\mathbf{v}\|_{L_2(\Omega)} \leq \|\mathbf{v}\|_{L_2(\Omega;\mathcal{T}_h)} \lesssim \|\mathbf{v}\|_{L_2(\Omega)} \quad \forall \mathbf{v} \in V_h.$$

Let Π_h be the nodal interpolation operator for the Raviart-Thomas-Nédélec finite element space V_h . It is well-known [37, 10] that

$$(2.3) \quad \|\boldsymbol{\zeta} - \Pi_h \boldsymbol{\zeta}\|_{L_2(\Omega)} \lesssim h^s |\boldsymbol{\zeta}|_{H^s(\Omega)} \quad \text{for } \frac{1}{2} < s \leq \ell + 1,$$

where $h = \max_{T \in \mathcal{T}_h} \text{diam } T$ is the mesh size. We also have, by a standard argument based on the Bramble-Hilbert lemma [11, 25],

$$(2.4) \quad \sum_{\sigma \in \mathfrak{S}_h} h_\sigma \|(\boldsymbol{\zeta} - \Pi_h \boldsymbol{\zeta}) \cdot \mathbf{n}_\sigma\|_{L_2(\sigma)}^2 \lesssim h^{2s} |\boldsymbol{\zeta}|_{H^s(\Omega)}^2 \quad \text{for } \frac{1}{2} < s \leq \ell + 1.$$

The norm $\|\cdot\|_{H^1(\Omega;\mathcal{T}_h)}$ on $H^1(\Omega) + Q_h$ is defined by

$$(2.5) \quad \|q\|_{H^1(\Omega;\mathcal{T}_h)}^2 = \sum_{T \in \mathcal{T}_h} \|\nabla q\|_{L_2(T)}^2 + \sum_{\sigma \in \mathfrak{S}_h} \frac{1}{h_\sigma} \|[[q]]_\sigma\|_{L_2(\sigma)}^2,$$

where $[[q]]_\sigma$ is the jump of $q \in Q_h$ across a side $\sigma \in \mathfrak{S}_h$ defined as follows.

If σ is interior to Ω , then σ is the common side of the elements T_\pm and

$$(2.6) \quad [[q]]_\sigma = q_- \mathbf{n}_{\sigma,-} + q_+ \mathbf{n}_{\sigma,+},$$

where $q_\pm = q|_{T_\pm}$ and $\mathbf{n}_{\sigma,\pm}$ is the unit normal of σ pointing towards the outside of T_\pm .

If σ is on $\partial\Omega$, then σ is the side of a unique element T_σ in \mathcal{T}_h and

$$(2.7) \quad [[q]]_\sigma = q_{T_\sigma} \mathbf{n}_{T_\sigma},$$

where $q_{T_\sigma} = q|_{T_\sigma}$ and \mathbf{n}_{T_σ} is the unit normal of σ pointing towards the outside of T_σ .

The norm $\|\cdot\|_{H^1(\Omega;\mathcal{T}_h)}$ is a well-known norm in the analysis of discontinuous Galerkin methods for second order problems [4, 18], and we have a standard interpolation error estimate

$$(2.8) \quad \|\phi - \mathcal{I}_h \phi\|_{H^1(\Omega;\mathcal{T}_h)} \lesssim h^s |\phi|_{H^{1+s}(\Omega)} \quad \text{for } 0 < s \leq \ell,$$

where \mathcal{I}_h is the nodal interpolation operator for the conforming P_ℓ Lagrange finite element space.

Remark 2.1. The estimate (2.8) implies

$$\lim_{h \downarrow 0} \inf_{q \in Q_h} \|p - q\|_{H^1(\Omega;\mathcal{T}_h)} = 0,$$

which is not true if $\ell = 0$. This is the reason why we only consider Raviart-Thomas-Nédélec finite element methods of order $\ell \geq 1$.

Remark 2.2. The connection between the DG norm $\|\cdot\|_{H^1(\Omega; \mathcal{T}_h)}$ and the Raviart-Thomas-Nédélec finite element method was exploited in [39] for the preconditioning of the saddle point problem (1.8).

2.2. Stability Estimates. Since \mathbf{A} is a smooth symmetric positive definite matrix on $\bar{\Omega}$, we have the obvious estimates

$$(2.9) \quad a(\mathbf{w}, \mathbf{v}) \lesssim \|\mathbf{w}\|_{L_2(\Omega)} \|\mathbf{v}\|_{L_2(\Omega)} \quad \forall \mathbf{v}, \mathbf{w} \in [L_2(\Omega)]^d,$$

$$(2.10) \quad a(\mathbf{v}, \mathbf{v}) \gtrsim \|\mathbf{v}\|_{L_2(\Omega)}^2 \gtrsim \|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)}^2 \quad \forall \mathbf{v} \in V_h.$$

Let α be the index of elliptic regularity that appears in (1.11). It follows from integration by parts and (2.6)–(2.7) that

$$(2.11) \quad \begin{aligned} -\langle \nabla \cdot \mathbf{v}, q \rangle_{H^{-1+\alpha}(\Omega) \times H^{1-\alpha}(\Omega)} &= \sum_{T \in \mathcal{T}_h} \left(- \int_{\partial T} (\mathbf{n}_T \cdot \mathbf{v}) q \, ds + \int_T \mathbf{v} \cdot \nabla q \, dx \right) \\ &= - \sum_{\sigma \in \mathfrak{S}_h} \int_{\sigma} \mathbf{v} \cdot \llbracket q \rrbracket_{\sigma} \, ds + \sum_{T \in \mathcal{T}_h} \int_T \mathbf{v} \cdot \nabla q \, dx \end{aligned}$$

for all $\mathbf{v} \in [H^\alpha(\Omega)]^d + V_h$ and $q \in H^1(\Omega) + Q_h$, and therefore

$$(2.12) \quad \begin{aligned} b(\mathbf{v}, q) &\leq \left(\sum_{\sigma \in \mathfrak{S}_h} h_{\sigma} \|\mathbf{v} \cdot \mathbf{n}_{\sigma}\|_{L_2(\sigma)}^2 \right)^{\frac{1}{2}} \left(\sum_{\sigma \in \mathfrak{S}_h} h_{\sigma}^{-1} \|\llbracket q \rrbracket_{\sigma}\|_{L_2(\sigma)}^2 \right)^{\frac{1}{2}} \\ &\quad + \left(\sum_{T \in \mathcal{T}_h} \|\mathbf{v}\|_{L_2(T)}^2 \right)^{\frac{1}{2}} \left(\sum_{T \in \mathcal{T}_h} \|\nabla q\|_{L_2(T)}^2 \right)^{\frac{1}{2}} \\ &\leq \|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} \|q\|_{H^1(\Omega; \mathcal{T}_h)} \end{aligned}$$

for all $\mathbf{v} \in [H^\alpha(\Omega)]^d + V_h$ and $q \in H^1(\Omega) + Q_h$.

Given any $q \in Q_h$ (a piecewise P_ℓ function), we define $\mathbf{v}_q \in V_h$ by

$$(2.13) \quad \mathbf{v}_q \cdot \mathbf{n}_{\sigma} = -\frac{1}{h_{\sigma}} \llbracket q \rrbracket_{\sigma} \cdot \mathbf{n}_{\sigma} \quad \forall \sigma \in \mathfrak{S}_h,$$

$$(2.14) \quad \Xi_{T, \ell-1} \mathbf{v}_q = \nabla q_T \quad \forall T \in \mathcal{T}_h,$$

where $\Xi_{T, \ell-1}$ is the orthogonal projection from $[L_2(T)]^d$ onto $[P_{\ell-1}(T)]^d$. It follows from (2.13), (2.14), the definition of the Raviart-Thomas-Nédélec element [37, 10] and scaling that

$$(2.15) \quad \begin{aligned} \|\mathbf{v}_q\|_{L_2(\Omega; \mathcal{T}_h)}^2 &\approx \sum_{T \in \mathcal{T}_h} \|\Xi_{T, \ell-1} \mathbf{v}_q\|_{L_2(T)}^2 + \sum_{\sigma \in \mathfrak{S}_h} h_{\sigma} \|\mathbf{v}_q \cdot \mathbf{n}_{\sigma}\|_{L_2(\sigma)}^2 \\ &= \sum_{T \in \mathcal{T}_h} \|\nabla q\|_{L_2(T)}^2 + \sum_{\sigma \in \mathfrak{S}_h} \frac{1}{h_{\sigma}} \|\llbracket q \rrbracket_{\sigma}\|_{L_2(\sigma)}^2 = \|q\|_{H^1(\Omega; \mathcal{T}_h)}^2. \end{aligned}$$

On the other hand (2.11), (2.13) and (2.14) imply

$$(2.16) \quad b(\mathbf{v}_q, q) = \|q\|_{H^1(\Omega; \mathcal{T}_h)}^2.$$

Combining (2.12), (2.15) and (2.16), we arrive at the inf-sup condition

$$(2.17) \quad \sup_{\mathbf{v} \in V_h} \frac{b(\mathbf{v}, q)}{\|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)}} \approx \|q\|_{H^1(\Omega; \mathcal{T}_h)} \quad \forall q \in Q_h.$$

It follows from (2.9), (2.10), (2.12), (2.17) and the saddle point theory [6, 20, 10] that

$$(2.18) \quad \begin{aligned} & \|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)} \\ & \approx \sup_{(\mathbf{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r))}{\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \quad \forall (\mathbf{v}, q) \in V_h \times Q_h. \end{aligned}$$

Note that (2.9), (2.12) and (1.7) imply

$$(2.19) \quad \mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r)) \lesssim (\|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) (\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)})$$

for all $\mathbf{v}, \mathbf{w} \in [H^\alpha(\Omega)]^d + V_h$ and $q, r \in H^1(\Omega) + Q_h$,

2.3. Error Estimates. Let α be the index of elliptic regularity in (1.11) and $F \in H^{-1+\alpha}(\Omega)$. According to Remark 1.1, the system (1.8) is well-defined and the solution $(\mathbf{u}, p) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ of (1.9)–(1.10) satisfies

$$\mathcal{B}((\mathbf{u}, p), (\mathbf{v}, q)) = F(q) \quad \forall (\mathbf{v}, q) \in V_h \times Q_h.$$

Consequently we have the Galerkin relation

$$(2.20) \quad \mathcal{B}((\mathbf{u}, p), (\mathbf{v}, q)) = \mathcal{B}((\mathbf{u}_h, p_h), (\mathbf{v}, q)) \quad \forall (\mathbf{v}, q) \in V_h \times Q_h.$$

Let $(\mathbf{v}, q) \in V_h \times Q_h$ be arbitrary. It follows from (2.18)–(2.20) that

$$\begin{aligned} & \|\mathbf{v} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|q - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \\ & \approx \sup_{(\mathbf{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}((\mathbf{v} - \mathbf{u}_h, q - p_h), (\mathbf{w}, r))}{\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \\ & = \sup_{(\mathbf{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}((\mathbf{v} - \mathbf{u}, q - p), (\mathbf{w}, r))}{\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \lesssim \|\mathbf{v} - \mathbf{u}\|_{L_2(\Omega; \mathcal{T}_h)} + \|q - p\|_{H^1(\Omega; \mathcal{T}_h)} \end{aligned}$$

and hence

$$\begin{aligned} & \|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \\ & \leq \|\mathbf{u} - \mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - q\|_{H^1(\Omega; \mathcal{T}_h)} + \|\mathbf{v} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|q - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \\ & \lesssim \|\mathbf{u} - \mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - q\|_{H^1(\Omega; \mathcal{T}_h)}, \end{aligned}$$

which then implies the quasi-optimal error estimate

$$(2.21) \quad \|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim \inf_{\mathbf{v} \in V_h} \|\mathbf{u} - \mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \inf_{q \in Q_h} \|p - q\|_{H^1(\Omega; \mathcal{T}_h)}.$$

Putting (1.11), (2.3), (2.4), (2.8) and (2.21) together, we have

$$(2.22) \quad \|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim h^\alpha \|F\|_{H^{-1+\alpha}(\Omega)},$$

and in the case where $p \in H^m(\Omega)$ for $m \leq \ell + 1$,

$$(2.23) \quad \|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} + \|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim h^{m-1} |p|_{H^m(\Omega)}.$$

Remark 2.3. The estimates (2.22) and (2.23) in the nonconforming energy norm $\|\cdot\|_{L_2(\Omega; \mathcal{T}_h)} + \|\cdot\|_{H^1(\Omega; \mathcal{T}_h)}$ are more informative than the standard error estimates for the Raviart-Thomas-Nédélec finite element methods in [38, 27, 10] since they provide approximations of the flux on the element interfaces. One can also recover the standard error estimates from (2.22)–(2.23).

3. MULTIGRID METHODS

We will introduce the multigrid methods for (1.8) in this section. The operators involved are defined with respect to a mesh-dependent inner product, and the smoothers for pre-smoothing and post-smoothing are defined in terms of a block-diagonal preconditioner.

3.1. Set-Up. Let \mathcal{T}_0 be an initial triangulation of Ω and the triangulations $\mathcal{T}_1, \mathcal{T}_2, \dots$ be obtained from \mathcal{T}_0 through uniform subdivisions. Since the Raviart-Thomas-Nédélec finite element pairs $V_k \times Q_k$ associated with \mathcal{T}_k are nested, we take the coarse-to-fine intergrid transfer operator $I_{k-1}^k : V_{k-1} \times Q_{k-1} \longrightarrow V_k \times Q_k$ to be the natural injection and define the Ritz projection operator $P_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ by

$$(3.1) \quad \mathcal{B}(P_k^{k-1}(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}((\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r))$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}$.

Let $(\cdot, \cdot)_k$ be a mesh-dependent inner product on V_k such that

$$(3.2) \quad (\mathbf{v}, \mathbf{v})_k \approx \|\mathbf{v}\|_{L_2(\Omega)}^2 \quad \forall \mathbf{v} \in V_k$$

and the nodal basis (vector) functions for V_k are orthogonal with respect to $(\cdot, \cdot)_k$. Similarly, let $((\cdot, \cdot))_k$ be a mesh-dependent inner product on Q_k such that

$$(3.3) \quad ((q, q))_k \approx \|q\|_{L_2(\Omega)}^2 \quad \forall q \in Q_k$$

and the nodal basis functions for Q_k are orthogonal with respect to $((\cdot, \cdot))_k$.

Remark 3.1. The inner products $(\cdot, \cdot)_k$ and $((\cdot, \cdot))_k$ are constructed by mass lumping.

The mesh-dependent inner product $[\cdot, \cdot]_k$ on $V_k \times Q_k$ is then defined by

$$(3.4) \quad [(\mathbf{v}, q), (\mathbf{w}, r)]_k = h_k^2 (\mathbf{v}, \mathbf{w})_k + ((q, r))_k,$$

where $h_k = \max_{T \in \mathcal{T}_k} \text{diam } T$ is the mesh size of \mathcal{T}_k . We take the fine-to-coarse intergrid transfer operator $I_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ to be the transpose of I_{k-1}^k with respect to the mesh-dependent inner products on $V_k \times Q_k$ and $V_{k-1} \times Q_{k-1}$, i.e.,

$$(3.5) \quad [I_k^{k-1}(\mathbf{v}, q), (\mathbf{w}, r)]_{k-1} = [(\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r)]_k$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}$.

Let the system operator $\mathbb{B}_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ be defined by

$$(3.6) \quad [\mathbb{B}_k(\mathbf{v}, q), (\mathbf{w}, r)]_k = \mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k.$$

Our goal is to develop multigrid algorithms for problems of the form

$$(3.7) \quad \mathbb{B}_k(\mathbf{v}, q) = (\mathbf{g}, z).$$

3.2. A Block-Diagonal Preconditioner. Let $L_k : Q_k \longrightarrow Q_k$ be an operator that is SPD with respect to $((\cdot, \cdot))_k$ and satisfies

$$(3.8) \quad ((L_k^{-1}q, q))_k \approx \|q\|_{H^1(\Omega; \mathcal{T}_h)}^2 \quad \forall q \in Q_k.$$

Then the preconditioner $\mathbb{S}_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ given by

$$(3.9) \quad \mathbb{S}_k(\mathbf{v}, q) = (h_k^2 \mathbf{v}, L_k q)$$

is SPD with respect to $[\cdot, \cdot]_k$ and we have

$$(3.10) \quad [\mathbb{S}_k^{-1}(\mathbf{v}, q), (\mathbf{v}, q)]_k \approx \|\mathbf{v}\|_{L_2(\Omega)}^2 + \|q\|_{H^1(\Omega; \mathcal{T}_k)}^2 \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

by (3.2)–(3.4) and (3.8).

Remark 3.2. The operator L_k can be constructed through multigrid [29, 19] or domain decomposition [28, 34, 2]

The following result connects the operators \mathbb{B}_k , \mathbb{S}_k and the nonconforming energy norm for $V_k \times Q_k$.

Lemma 3.3. *The norm equivalence*

$$(3.11) \quad \left[\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k(\mathbf{v}, q), (\mathbf{v}, q) \right]_k^{\frac{1}{2}} \approx \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

holds for $k = 0, 1, 2, \dots$

Proof. Let $(\mathbf{v}, q) \in V_k \times Q_k$ be arbitrary and $(\mathbf{x}, y) = \mathbb{S}_k \mathbb{B}_k(\mathbf{v}, q)$. It follows from (2.2), (2.18), (3.6), (3.10) and duality that

$$\begin{aligned} [\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k(\mathbf{v}, q), (\mathbf{v}, q)]_k^{\frac{1}{2}} &= [\mathbb{S}_k^{-1}(\mathbb{S}_k \mathbb{B}_k)(\mathbf{v}, q), (\mathbb{S}_k \mathbb{B}_k)(\mathbf{v}, q)]_k^{\frac{1}{2}} \\ &= [\mathbb{S}_k^{-1}(\mathbf{x}, y), (\mathbf{x}, y)]_k^{\frac{1}{2}} \\ &= \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{[\mathbb{S}_k^{-1}(\mathbf{x}, y), (\mathbf{w}, r)]_k}{[\mathbb{S}_k^{-1}(\mathbf{w}, r), (\mathbf{w}, r)]_k^{\frac{1}{2}}} \\ &\approx \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{[\mathbb{B}_k(\mathbf{v}, q), (\mathbf{w}, r)]_k}{\|\mathbf{w}\|_{L_2(\Omega)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \\ &= \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r))}{\|\mathbf{w}\|_{L_2(\Omega)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \approx \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}. \end{aligned}$$

□

Let $\rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k)$ be the spectral radius of the operator $\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k$. It follows from (3.4), (3.11) and a standard inverse estimate [21, 18] that

$$(3.12) \quad \rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k) \lesssim h_k^{-2}.$$

We can therefore choose a damping factor δ_k of the form Ch_k^2 such that

$$(3.13) \quad \delta_k \cdot \rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k) \leq 1.$$

3.3. Multigrid Algorithms. Let the output of the W -cycle algorithm for (3.7) with initial guess (\mathbf{v}_0, q_0) and m_1 (resp. m_2) pre-smoothing (resp. post-smoothing) steps be denoted by $MG_W(k, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2)$.

We use a direct solve for $k = 0$, i.e., we take $MG_W(0, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2)$ to be $\mathbb{B}_0^{-1}(\mathbf{g}, z)$. For $k \geq 1$, we compute $MG_W(k, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2)$ in three steps.

Pre-Smoothing The approximate solutions $(\mathbf{v}_1, q_1), \dots, (\mathbf{v}_{m_1}, q_{m_1})$ are computed recursively by

$$(3.14) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{S}_k \mathbb{B}_k((\mathbf{g}, z) - \mathbb{B}_k(\mathbf{v}_{j-1}, q_{j-1}))$$

for $1 \leq j \leq m_1$, where the damping factor δ_k satisfies (3.13).

Coarse Grid Correction Let $(\mathbf{g}', z') = I_k^{k-1}((\mathbf{g}, z) - \mathbb{B}_k(\mathbf{v}_{m_1}, q_{m_1}))$ be the transferred residual of $(\mathbf{v}_{m_1}, q_{m_1})$ and compute $(\mathbf{v}'_1, q'_1), (\mathbf{v}'_2, q'_2) \in V_{k-1} \times Q_{k-1}$ by

$$(3.15) \quad (\mathbf{v}'_1, q'_1) = MG_W(k-1, (\mathbf{g}', z'), (\mathbf{0}, 0), m_1, m_2),$$

$$(3.16) \quad (\mathbf{v}'_2, q'_2) = MG_W(k-1, (\mathbf{g}', z'), (\mathbf{v}'_1, q'_1), m_1, m_2).$$

We then take $(\mathbf{v}_{m_1+1}, q_{m_1+1})$ to be $(\mathbf{v}_{m_1}, q_{m_1}) + I_{k-1}^k(\mathbf{v}'_2, q'_2)$.

Post-Smoothing The approximate solutions $(\mathbf{v}_{m_1+1}, q_{m_1+1}), \dots, (\mathbf{v}_{m_1+m_2+1}, q_{m_1+m_2+1})$ are computed recursively by

$$(3.17) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{B}_k \mathbb{S}_k((\mathbf{g}, z) - \mathbb{B}_k(\mathbf{v}_{j-1}, q_{j-1}))$$

for $m_1 + 2 \leq j \leq m_1 + m_2 + 1$.

The final output is $MG_W(k, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2) = (\mathbf{v}_{m_1+m_2+1}, q_{m_1+m_2+1})$.

Let $MG_V(k, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2)$ be the output of the V -cycle algorithm for (3.7) with initial guess (\mathbf{v}_0, q_0) and m_1 (resp. m_2) pre-smoothing (resp. post-smoothing) steps. The computation of $MG_V(k, (\mathbf{g}, z), (\mathbf{v}_0, q_0), m_1, m_2)$ differs from the computation for the W -cycle algorithm only in the coarse grid correction step, where we compute

$$(\mathbf{v}'_1, q'_1) = MG_V(k-1, (\mathbf{g}', z'), (\mathbf{0}, 0), m_1, m_2)$$

and take $(\mathbf{v}_{m_1+1}, q_{m_1+1})$ to be $(\mathbf{v}_{m_1}, q_{m_1}) + I_{k-1}^k(\mathbf{v}'_1, q'_1)$.

3.4. Error Propagation Operators. The effect of one post-smoothing step defined by (3.17) is measured by

$$(3.18) \quad R_k = Id_k - \delta_k \mathbb{B}_k \mathbb{S}_k \mathbb{B}_k,$$

where $Id_k : V_k \times Q_k \rightarrow V_k \times Q_k$ is the identity operator. The choice of the smoother $\mathbb{B}_k \mathbb{S}_k$ for post-smoothing is motivated by the fact that (3.18) is the error propagation operator of one Richardson relaxation step for the SPD problem

$$(3.19) \quad \mathbb{B}_k \mathbb{S}_k \mathbb{B}_k(\mathbf{v}, q) = \mathbb{B}_k \mathbb{S}_k(\mathbf{g}, z),$$

which is equivalent to (3.7).

On the other hand, the effect of one pre-smoothing step defined by (3.14) is measured by

$$(3.20) \quad S_k = Id_k - \delta_k \mathbb{S}_k \mathbb{B}_k^2.$$

Our choice of the smoother $\mathbb{S}_k \mathbb{B}_k$ for the pre-smoothing is motivated by the adjoint relation

$$(3.21) \quad \mathcal{B}(R_k(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}((\mathbf{v}, q), S_k(\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k$$

that follows from (3.6), (3.18) and (3.20).

The error propagation operator $E_k : V_k \times Q_k \rightarrow V_k \times Q_k$ for the multigrid algorithms satisfies the well-known recursive relation [31, 13, 18]

$$(3.22) \quad E_k = R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1} + I_{k-1}^k E_{k-1}^p P_k^{k-1}) S_k^{m_1} \quad \text{for } k \geq 1,$$

where P_k^{k-1} is the Ritz projection operator defined in (3.1) and $p = 2$ (resp. 1) for the W -cycle (resp. V -cycle) algorithm.

Since I_{k-1}^k is the natural injection, we have

$$(3.23) \quad P_k^{k-1} I_{k-1}^k = Id_{k-1}, \quad (Id_k - I_{k-1}^k P_k^{k-1})^2 = Id_k - I_{k-1}^k P_k^{k-1},$$

and the Galerkin orthogonality

$$(3.24) \quad 0 = \mathcal{B}((Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r))$$

that is valid for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}$.

4. MESH-DEPENDENT NORMS FOR MULTIGRID ANALYSIS

We introduce in this section a scale of mesh-dependent norms that are crucial for the convergence analysis of the W -cycle multigrid algorithm in Section 5.

4.1. Definition of the Mesh-Dependent Norms. For $0 \leq s \leq 1$, we define the scale of mesh-dependent norms $\|\cdot\|_{s,k}$ in terms of the SPD operator $\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k$ and the mesh-dependent inner product $[\cdot, \cdot]_k$ as follows:

$$(4.1) \quad \|(\mathbf{v}, q)\|_{s,k}^2 = [(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k)^s(\mathbf{v}, q), (\mathbf{v}, q)]_k \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

In view of (3.2)–(3.4), (3.11) and (4.1), we have the obvious norm equivalences

$$(4.2) \quad \|(\mathbf{v}, q)\|_{0,k}^2 \approx h_k^2 \|\mathbf{v}\|_{L_2(\Omega)}^2 + \|q\|_{L_2(\Omega)}^2 \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

$$(4.3) \quad \|(\mathbf{v}, q)\|_{1,k}^2 \approx \|\mathbf{v}\|_{L_2(\Omega)}^2 + \|q\|_{H^1(\Omega; \mathcal{T}_k)}^2 \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

Thus the $\|\cdot\|_{1,k}$ norm is equivalent to the nonconforming energy norm on $V_k \times Q_k$ and we have the following stability result.

Lemma 4.1. *The operators I_{k-1}^k and P_k^{k-1} are stable with respect to the mesh-dependent norm $\|\cdot\|_{1,k}$.*

Proof. Since I_{k-1}^k is the natural injection, the stability estimate

$$\begin{aligned} \|I_{k-1}^k(\mathbf{w}, r)\|_{1,k} &\approx \|\mathbf{w}\|_{L_2(\Omega)} + \|r\|_{H^1(\Omega; \mathcal{T}_k)} \\ &\lesssim \|\mathbf{w}\|_{L_2(\Omega)} + \|r\|_{H^1(\Omega; \mathcal{T}_{k-1})} \approx \|(\mathbf{w}, r)\|_{1,k-1} \quad \forall (\mathbf{w}, r) \in V_{k-1} \times Q_{k-1} \end{aligned}$$

follows from (2.5), (4.3) and a direct calculation.

The stability of P_k^{k-1} then follows from (2.2), (2.18), (2.19), (3.1), (4.3) and duality:

$$\begin{aligned} \|P_k^{k-1}(\mathbf{v}, q)\|_{1,k-1} &\approx \sup_{(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}} \frac{\mathcal{B}(P_k^{k-1}(\mathbf{v}, q), (\mathbf{w}, r))}{\|(\mathbf{w}, r)\|_{1,k-1}} \\ &= \sup_{(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}} \frac{\mathcal{B}((\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r))}{\|(\mathbf{w}, r)\|_{1,k-1}} \\ &\lesssim \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k. \end{aligned}$$

□

We will need a connection between $\|\cdot\|_{1-\alpha,k}$ and a Sobolev norm in the proof of the approximation property in Section 5. Towards this goal we introduce the operator $D_k : Q_k \rightarrow Q_k$ defined by

$$(4.4) \quad ((D_k q, r))_k = \sum_{T \in \mathcal{T}_k} \int_T \nabla q \cdot \nabla r \, dx + \sum_{\sigma \in \mathfrak{S}_k} h_\sigma^{-1} \int_\sigma [[q]]_\sigma \cdot [[r]]_\sigma \, ds \quad \forall (q, r) \in V_k \times Q_k.$$

Then D_k is SPD with respect to $((\cdot, \cdot))_k$ and the relations

$$(4.5) \quad ((D_k^0 q, q))_k \approx \|q\|_{L_2(\Omega)}^2 \quad \forall q \in Q_k,$$

$$(4.6) \quad ((D_k q, q))_k \approx \|q\|_{H^1(\Omega; \mathcal{T}_k)}^2 \quad \forall q \in Q_k,$$

follow immediately from (2.5), (3.3) and (4.4).

Remark 4.2. The operator L_k that appears in (3.9) is just an optimal preconditioner of D_k .

It follows from standard inverse estimates that $\rho(D_k) \lesssim h_k^{-2}$ and hence we have, by the spectral theorem,

$$(4.7) \quad \|q\|_{H^1(\Omega; \mathcal{T}_k)}^2 = ((D_k q, q))_k \leq C h_k^{2(s-1)} ((D_k^s q, q))_k \quad \forall q \in Q_k.$$

In view of (4.2), (4.3), (4.5) and (4.6), we have

$$\begin{aligned} \|(\mathbf{v}, q)\|_{0,k}^2 &\approx h_k^2 \|\mathbf{v}\|_{L_2(\Omega)}^2 + ((D_k^0 q, q))_k \quad \forall q \in Q_k, \\ \|(\mathbf{v}, q)\|_{1,k}^2 &\approx \|\mathbf{v}\|_{L_2(\Omega)}^2 + ((D_k^1 q, q))_k \quad \forall q \in Q_k, \end{aligned}$$

which imply, through interpolation between Hilbert scales [43, Chapter 23], the norm equivalence

$$(4.8) \quad \|(\mathbf{v}, q)\|_{s,k}^2 \approx h_k^{2(1-s)} \|\mathbf{v}\|_{L_2(\Omega)}^2 + ((D_k^s q, q))_k \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

that holds for $0 \leq s \leq 1$.

It only remains to relate $((D_k^s q, q))_k$ to Sobolev norms, which will require certain tools from the multigrid theory for nonconforming finite element methods [15, 19].

4.2. Enriching and Forgetting Operators. Let $\tilde{Q}_k \subset H^1(\Omega)$ be the $P_{d+1+\ell}$ Lagrange finite element space associated with \mathcal{T}_k . The *enriching* operator $\mathcal{E}_k : Q_k \rightarrow \tilde{Q}_k$ is defined by averaging, i.e.,

$$(4.9) \quad (\mathcal{E}_k q)(x) = \frac{1}{|\mathcal{T}_x|} \sum_{T \in \mathcal{T}_x} q_T(x),$$

where x is any node for \tilde{Q}_k , \mathcal{T}_x is the set of the elements in \mathcal{T}_k that share the node x , and $|\mathcal{T}_x|$ is the number of elements in \mathcal{T}_x .

The following estimate is obtained by a straight-forward local calculation:

$$(4.10) \quad \sum_{T \in \mathcal{T}_k} h_T^{-2} \|q - \mathcal{E}_k q\|_{L_2(T)}^2 \lesssim \sum_{\sigma \in \mathfrak{S}_k^i} \frac{1}{h_\sigma} \|\llbracket q \rrbracket_\sigma\|_{L_2(\sigma)}^2 \quad \forall q \in Q_k,$$

where $h_T = \text{diam } T$ and \mathfrak{S}_k^i is the set of the interior faces.

Since q and $\mathcal{E}_k q$ agree at the $(\ell+1)(\ell+2)/2$ interior nodes for each $T \in \mathcal{T}_k$ when $d=2$ and the $(\ell+1)(\ell+2)(\ell+3)/6$ interior nodes for each $T \in \mathcal{T}_k$ when $d=3$, we can define a *forgetting* operator $\mathcal{F}_k : \tilde{Q}_k \rightarrow Q_k$ element by element so that

$$(4.11) \quad \mathcal{F}_k \circ \mathcal{E}_k = Id_k$$

as follows. For any $\tilde{q} \in \tilde{Q}_k$, we define $\mathcal{F}_k \tilde{q}$ to be the (unique) function $q \in Q_k$ such that, for any $T \in \mathcal{T}_k$, $q = \tilde{q}$ at the nodes of \tilde{Q}_k interior to T . We have, by scaling,

$$(4.12) \quad \|\tilde{q} - \mathcal{F}_k \tilde{q}\|_{L_2(T)} \leq Ch_T |\tilde{q}|_{H^1(T)} \quad \forall \tilde{q} \in \tilde{Q}_k, T \in \mathcal{T}_k.$$

The estimates (4.10) and (4.12) then imply, through standard inverse estimates [21, 18],

$$(4.13) \quad \|\mathcal{E}_k q\|_{H^1(\Omega)} \leq C \|q\|_{H^1(\Omega; \mathcal{T}_k)} \quad \forall q \in Q_k,$$

$$(4.14) \quad \|\mathcal{E}_k q\|_{L_2(\Omega)} \leq C \|q\|_{L_2(\Omega)} \quad \forall q \in Q_k,$$

$$(4.15) \quad \|q - \mathcal{E}_k q\|_{L_2(\Omega)} \leq C_s h_k^s \|q\|_{H^s(\Omega)} \quad \forall q \in Q_k, 0 \leq s < \frac{1}{2},$$

$$(4.16) \quad \|\mathcal{F}_k \tilde{q}\|_{H^1(\Omega; \mathcal{T}_k)} \leq C \|\tilde{q}\|_{H^1(\Omega)} \quad \forall \tilde{q} \in \tilde{Q}_k,$$

$$(4.17) \quad \|\mathcal{F}_k \tilde{q}\|_{L_2(\Omega)} \leq C \|\tilde{q}\|_{L_2(\Omega)} \quad \forall \tilde{q} \in \tilde{Q}_k.$$

4.3. Equivalence between Mesh-Dependent Norms and Sobolev Norms. We will connect the mesh-dependent norms $\|\cdot\|_{s,k}$ to the Sobolev norms through two lemmas.

Lemma 4.3. *The norm equivalence*

$$((D_k^s q, q))_k \approx \|\mathcal{E}_k q\|_{H^s(\Omega)}^2 \quad \forall q \in Q_k$$

holds for $0 \leq s \leq 1$.

Proof. It follows from the estimates (4.5), (4.6), (4.13), (4.14) and interpolation between Hilbert scales that

$$\|\mathcal{E}_k q\|_{H^s(\Omega)} \lesssim ((D_k^s q, q))_k^{\frac{1}{2}} \quad \forall q \in Q_k.$$

In order to prove the estimate in the opposite direction, we introduce the operator

$$J_k = \mathcal{F}_k \circ \Lambda_k,$$

where $\Lambda_k : L_2(\Omega) \rightarrow \tilde{Q}_k$ is the orthogonal projection. In view of (4.11), we have

$$(4.18) \quad J_k \mathcal{E}_k q = \mathcal{F}_k \Lambda_k \mathcal{E}_k q = \mathcal{F}_k \mathcal{E}_k q = q \quad \forall q \in Q_k.$$

Moreover, it follows from (4.5), (4.6), (4.16), (4.17) and the well-known estimate [12]

$$\|\Lambda_k \zeta\|_{H^1(\Omega)} \lesssim \|\zeta\|_{H^1(\Omega)} \quad \forall \zeta \in H^1(\Omega)$$

that

$$\begin{aligned} ((D_k^0 J_k \zeta, J_k \zeta))_k^{\frac{1}{2}} &\approx \|J_k \zeta\|_{L_2(\Omega)} \lesssim \|\zeta\|_{L_2(\Omega)} \quad \forall \zeta \in L_2(\Omega), \\ ((D_k^1 J_k \zeta, J_k \zeta))_k^{\frac{1}{2}} &\approx \|J_k \zeta\|_{H^1(\Omega; \mathcal{T}_k)} \lesssim \|\zeta\|_{H^1(\Omega)} \quad \forall \zeta \in H^1(\Omega). \end{aligned}$$

The two last estimates imply, by interpolation between Hilbert scales,

$$((D_k^s J_k \zeta, J_k \zeta))_k^{\frac{1}{2}} \leq C \|\zeta\|_{H^s(\Omega)} \quad \forall \zeta \in H^s(\Omega),$$

and hence, because of (4.18),

$$((D_k^s q, q))_k^{\frac{1}{2}} = ((D_k^s J_k \mathcal{E}_k q, J_k \mathcal{E}_k q))_k^{\frac{1}{2}} \lesssim \|\mathcal{E}_k q\|_{H^s(\Omega)} \quad \forall q \in Q_k.$$

□

Lemma 4.4. *For any $s \in [0, \frac{1}{2})$, we have*

$$((D_k^s q, q))_k \approx |q|_{H^s(\Omega)}^2 \quad \forall q \in Q_k,$$

where the constants in the norm equivalence depend on s .

Proof. Using the (non-standard) inverse estimate [8]

$$(4.19) \quad \|q - \mathcal{E}_k q\|_{H^s(\Omega)} \leq C_s h_k^{-s} \|q - \mathcal{E}_k q\|_{L_2(\Omega)} \quad \forall q \in Q_k \quad \text{and} \quad 0 \leq s < \frac{1}{2}$$

together with (4.7), (4.10) and Lemma 4.3, we find

$$\begin{aligned} \|q\|_{H^s(\Omega)} &\leq \|q - \mathcal{E}_k q\|_{H^s(\Omega)} + \|\mathcal{E}_k q\|_{H^s(\Omega)} \\ &\lesssim h_k^{-s} \|q - \mathcal{E}_k q\|_{L_2(\Omega)} + ((D_k^s q, q))_k^{\frac{1}{2}} \end{aligned}$$

$$\lesssim h_k^{1-s} \|q\|_{H^1(\Omega; \mathcal{T}_k)} + ((D_k^s q, q))_k^{\frac{1}{2}} \lesssim ((D_k^s q, q))_k^{\frac{1}{2}} \quad \forall q \in Q_k.$$

In the other direction we have, by (4.15), Lemma 4.3 and (4.19),

$$\begin{aligned} ((D_k^s q, q))_k^{\frac{1}{2}} &\lesssim \|\mathcal{E}_k q\|_{H^s(\Omega)} \\ &\lesssim (\|q\|_{H^s(\Omega)} + \|q - \mathcal{E}_k q\|_{H^s(\Omega)}) \\ &\lesssim (\|q\|_{H^s(\Omega)} + h_k^{-s} \|q - \mathcal{E}_k q\|_{L_2(\Omega)}) \lesssim \|q\|_{H^s(\Omega)} \quad \forall q \in Q_k. \end{aligned}$$

□

Combining (4.8) and Lemma 4.4, we arrive at the following result.

Corollary 4.5. *For any $s \in [0, \frac{1}{2})$, we have*

$$\|(\mathbf{v}, q)\|_{s,k} \approx h_k^{1-s} \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^s(\Omega)} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

where the constants in the norm equivalence depend on s .

4.4. Another Scale of Mesh-Dependent Norms. In Section 5. we will use the scale of mesh-dependent norms $\|\cdot\|_{s,k}$ to analyze the effect of post-smoothing coupled with coarse grid correction. In order to analyze the effect of pre-smoothing coupled with coarse grid correction, we will need a second scale of mesh-dependent norms $\|\!\!\| \cdot \|\!\!\|_{s,k}$.

For $1 \leq s \leq 2$, we define the mesh-dependent norm $\|\!\!\| \cdot \|\!\!\|_{s,k}$ by duality:

$$(4.20) \quad \|\!\!\|(\mathbf{v}, q)\|\!\!\|_{s,k} = \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r))}{\|\!\!\|(\mathbf{w}, r)\|\!\!\|_{2-s,k}} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

It follows from (2.18), (4.3) and (4.20) that

$$(4.21) \quad \|\!\!\|(\mathbf{v}, q)\|\!\!\|_{1,k} \approx \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_k)} \approx \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

Note that the two scales of mesh-dependent norms together provide a generalized Cauchy-Schwarz inequality for the bilinear form $\mathcal{B}(\cdot, \cdot)$:

$$(4.22) \quad \mathcal{B}((\mathbf{v}, q), (\mathbf{w}, r)) \leq \|\!\!\|(\mathbf{v}, q)\|\!\!\|_{1+\tau,k} \|\!\!\|(\mathbf{w}, r)\|\!\!\|_{1-\tau,k}$$

for all $(\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k$ and $0 \leq \tau \leq 1$.

5. CONVERGENCE ANALYSIS

In this section we will carry out the convergence analysis for the W -cycle algorithm, which is based on the smoothing and approximation properties [7, 31] with respect to the mesh-dependent norms in Section 4. Once we have established these properties with respect to the scale of mesh-dependent norms defined in Section 4.1, the analysis will proceed as in [17, Section 5.3].

Numerical results indicate that the V -cycle algorithm is also uniformly convergent in the nonconforming energy norm. But we will not consider the much more involved convergence analysis of the V -cycle algorithm in this paper.

5.1. Smoothing and Approximation Properties. Since the post-smoothing step in (3.17) is just the Richardson relaxation for the SPD problem (3.19) and the operator $\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k$ behaves like a typical SPD operator for second order problems (cf. (3.12)), we have a standard smoothing property whose proof is identical to that of [17, Lemma 5.1].

Lemma 5.1. *The estimate*

$$(5.1) \quad \|R_k^m(\mathbf{v}, q)\|_{1,k} \lesssim h_k^{-\tau} m^{-\tau/2} \|(\mathbf{v}, q)\|_{1-\tau,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

holds for $\tau \in [0, 1]$.

The following approximation property is based on Corollary 4.5 and a duality argument.

Lemma 5.2. *We have*

$$\|(Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)\|_{1-\alpha,k} \lesssim h_k^\alpha \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

where $\alpha \in (\frac{1}{2}, 1]$ is the index of elliptic regularity that appears in (1.11).

Proof. Let $(\mathbf{v}, q) \in V_k \times Q_k$ be arbitrary and $(\boldsymbol{\zeta}, \mu) = (Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)$. In view of Corollary 4.5, it suffices to show that

$$(5.2) \quad h_k^\alpha \|\boldsymbol{\zeta}\|_{L_2(\Omega)} + \|\mu\|_{H^{1-\alpha}(\Omega)} \lesssim h^\alpha \|(\mathbf{v}, q)\|_{1,k}.$$

The estimate for $\boldsymbol{\zeta}$ follows immediately from (4.3) and Lemma 4.1:

$$(5.3) \quad h_k^\alpha \|\boldsymbol{\zeta}\|_{L_2(\Omega)} \lesssim h_k^\alpha \|(\boldsymbol{\zeta}, \mu)\|_{1,k} \lesssim h^\alpha \|(\mathbf{v}, q)\|_{1,k}.$$

The estimate for μ is established through a duality argument. Let $\phi \in H^{-1+\alpha}(\Omega)$ and $(\boldsymbol{\xi}, \theta) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ satisfy (1.9)–(1.10) with F replaced by ϕ . Then we have

$$(5.4) \quad \mathcal{B}((\boldsymbol{\xi}, \theta), (\mathbf{w}, r)) = \phi(r) \quad \forall (\mathbf{w}, r) \in V_k \times Q_k$$

by Remark 1.1. Moreover, if we define $(\boldsymbol{\xi}_{k-1}, \theta_{k-1}) \in V_{k-1} \times Q_{k-1}$ by

$$(5.5) \quad \mathcal{B}((\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (\mathbf{w}, r)) = \phi(r) \quad \forall (\mathbf{w}, r) \in V_{k-1} \times Q_{k-1},$$

then

$$(5.6) \quad \|\boldsymbol{\xi} - \boldsymbol{\xi}_{k-1}\|_{L_2(\Omega; \mathcal{T}_k)} + \|\theta - \theta_{k-1}\|_{H^1(\Omega; \mathcal{T}_k)} \lesssim h_k^\alpha \|\phi\|_{H^{-1+\alpha}(\Omega)}$$

by the discretization error estimate (2.22), since $h_{k-1} \approx h_k$.

It follows from (2.2), (2.19), (3.24), (4.3) and (5.5)–(5.6) that

$$\begin{aligned} \phi(\mu) &= \mathcal{B}((\boldsymbol{\xi}, \theta), (\boldsymbol{\zeta}, \mu)) \\ &= \mathcal{B}((\boldsymbol{\xi}, \theta), (Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)) \\ &= \mathcal{B}((\boldsymbol{\xi}, \theta) - (\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)) \\ &= \mathcal{B}((\boldsymbol{\xi}, \theta) - (\boldsymbol{\xi}_{k-1}, \theta_{k-1}), (\mathbf{v}, q)) \\ &\lesssim (\|\boldsymbol{\xi} - \boldsymbol{\xi}_{k-1}\|_{L_2(\Omega; \mathcal{T}_k)} + \|\theta - \theta_{k-1}\|_{H^1(\Omega; \mathcal{T}_k)}) (\|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_k)} + \|q\|_{H^1(\Omega; \mathcal{T}_k)}) \\ &\lesssim h_k^\alpha \|\phi\|_{H^{-1+\alpha}(\Omega)} \|(\mathbf{v}, q)\|_{1,k} \end{aligned}$$

and hence, by duality,

$$(5.7) \quad \|\mu\|_{H^{1-\alpha}(\Omega)} = \sup_{\phi \in H^{-1+\alpha}(\Omega)} \frac{\phi(\mu)}{\|\phi\|_{H^{-1+\alpha}(\Omega)}} \lesssim h_k^\alpha \|(\mathbf{v}, q)\|_{1,k}.$$

The estimate (5.2) follows from (5.3) and (5.7). \square

5.2. Convergence of the Two-Grid Algorithm. In the two-grid algorithm the coarse grid residual equation is solved exactly. We can therefore set $E_{k-1} = 0$ in (3.22) to obtain the error propagation of the two-grid algorithm, which is given by $R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}$.

For $m \geq 1$, we have the following estimate on the effect of post-smoothing coupled with coarse grid correction by combining Lemma 5.1 and Lemma 5.2.

$$(5.8) \quad \|R_k^m(Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)\|_{1,k} \lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

Using (2.18), (3.1), (3.21), (4.20), (4.21) and (5.8), we then obtain the following estimate on the effect of pre-smoothing coupled with coarse grid correction, where $m \geq 1$.

$$(5.9) \quad \|(Id_k - I_{k-1}^k P_k^{k-1})S_k^m(\mathbf{v}, q)\|_{1,k} \lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

Therefore, for $m_1, m_2 \geq 1$, we have

$$(5.10) \quad \begin{aligned} & \|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\mathbf{v}, q)\|_{1,k} \\ &= \|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\mathbf{v}, q)\|_{1,k} \\ &\lesssim (m_1 m_2)^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k \end{aligned}$$

by (3.23), (4.21), (5.8) and (5.9).

Remark 5.3. Since the estimates (5.8)–(5.10) are identical to the estimates (5.7)–(5.9) in [17], we keep the arguments brief here and refer to [17, Section 5] for the details.

Putting (5.8)–(5.10) together, we arrive at the estimate

$$(5.11) \quad \|R_k^{m_2}(Id_k - I_{k-1}^k P_k^{k-1})S_k^{m_1}(\mathbf{v}, q)\|_{1,k} \leq C_* [\max(1, m_1) \max(m_2, 1)]^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $k \geq 1$. Thus the two-grid algorithm is a contraction if $\max(1, m_1) \max(m_2, 1)$ is sufficiently large.

5.3. Convergence of the W-Cycle Algorithm. The estimate (5.11) and a perturbation argument lead to the following result for the W -cycle algorithm, whose proof is identical to that of [17, Theorem 5.5].

Theorem 5.4. *Let E_k be the error propagation operator for the k -th level W -cycle algorithm. For any $C_\dagger > C_*$ (the constant in (5.11)), there exists a positive number m_* (independent of k) such that*

$$\|E_k(\mathbf{v}, q)\|_{1,k} \leq C_\dagger (\max(1, m_1) \max(1, m_2))^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $k \geq 1$, provided $\max(1, m_1) \max(1, m_2) \geq m_*$.

Therefore, if $\max(1, m_1) \max(1, m_2)$ (independent of k) is sufficiently large, then the W -cycle algorithm is a contraction with respect to the nonconforming energy norm and the contraction number is bounded away from 1 for $k \geq 1$, i.e., the W -cycle algorithm converges uniformly.

6. GENERAL SECOND ORDER ELLIPTIC PROBLEMS

In this section we extend the multigrid results for the Darcy system to general second order elliptic problems of the form

$$(6.1) \quad -\nabla \cdot (\mathbf{A} \nabla p) + \boldsymbol{\beta} \cdot \nabla p + \gamma p = f \quad \text{in } \Omega \quad \text{and} \quad p = g \quad \text{on } \partial\Omega,$$

which include the Darcy system (1.1)–(1.3) as a special case, together with the adjoint problems

$$(6.2) \quad -\nabla \cdot (\mathbf{A} \nabla p) - \nabla \cdot (\boldsymbol{\beta} p) + \gamma p = f \quad \text{in } \Omega \quad \text{and} \quad p = g \quad \text{on } \partial\Omega.$$

For the design and analysis of multigrid methods, it suffices to consider the case where $g = 0$. We assume that \mathbf{A} is a (sufficiently) smooth SPD $d \times d$ matrix function on $\bar{\Omega}$, $\boldsymbol{\beta} \in [W_\infty^1(\Omega)]^d$ and $\gamma \in L_\infty(\Omega)$. We also assume that the boundary value problems (6.1) and (6.2) are both well-posed, which is the case if, for example,

$$(6.3) \quad \gamma - \frac{1}{2} \nabla \cdot \boldsymbol{\beta} \geq 0 \quad \text{a.e. in } \Omega.$$

6.1. Finite Element Methods. The mixed finite element method for (6.1) is to find $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ such that

$$(6.4) \quad a(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) = 0 \quad \forall \mathbf{v} \in V_h,$$

$$(6.5) \quad b(\mathbf{u}_h, q) - c_h(p_h, q) = F(q) \quad \forall q \in Q_h,$$

where the finite element space $V_h \times Q_h$, the bilinear forms $a(\cdot, \cdot)$, $b(\cdot, \cdot)$ and the bounded linear functional F are identical to the ones for the Darcy system, and the mesh-dependent bilinear form $c_h(\cdot, \cdot)$ is defined by

$$(6.6) \quad c_h(r, q) = \int_{\Omega} (\gamma r + \boldsymbol{\beta} \cdot \nabla_h r) q \, dx \quad \forall q, r \in Q_h.$$

Here ∇_h is the piecewise defined gradient operator.

We will treat (6.4)–(6.5) as a nonconforming method for the following weak formulation of (6.1): Find $(\mathbf{u}, p) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ such that

$$(6.7) \quad a(\mathbf{u}, \mathbf{v}) + b'(\mathbf{v}, p) = 0 \quad \forall \mathbf{v} \in [L_2(\Omega)]^d,$$

$$(6.8) \quad b'(\mathbf{u}, q) - c(p, q) = F(q) \quad \forall q \in H_0^1(\Omega),$$

where the bilinear form $b'(\cdot, \cdot)$ is identical to the one in (1.9)–(1.10) and

$$c(r, q) = \int_{\Omega} (\gamma r + \boldsymbol{\beta} \cdot \nabla r) q \, dx.$$

Similarly, the mixed finite element method for the adjoint problem (6.2) is to find $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ such that

$$(6.9) \quad a(\mathbf{u}_h, \mathbf{v}) + b(\mathbf{v}, p_h) = 0 \quad \forall \mathbf{v} \in V_h,$$

$$(6.10) \quad b(\mathbf{u}_h, q) - c_h(q, p_h) = F(q) \quad \forall q \in Q_h,$$

and it can be treated as a nonconforming method for the following weak formulation of (6.2): Find $(\mathbf{u}, p) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ such that

$$(6.11) \quad a(\mathbf{u}, \mathbf{v}) + b'(\mathbf{v}, p) = 0 \quad \forall \mathbf{v} \in [L_2(\Omega)]^d,$$

$$(6.12) \quad b'(\mathbf{u}, q) - c(q, p) = F(q) \quad \forall q \in H_0^1(\Omega).$$

Remark 6.1. The discretizations (6.4)–(6.5) and (6.9)–(6.10) for the convection-diffusion-reaction problem (6.1) and the advection-diffusion-reaction problem (6.2) are different from the mixed finite element methods in [24] which are based on $H(\text{div}; \Omega) \times L_2(\Omega)$ formulations. Instead, they are related to the upwind mixed finite element methods in [32].

Remark 6.2. Note that the systems (6.7)–(6.8) and (6.11)–(6.12) are well-posed for $F \in H^{-s}(\Omega)$ ($0 \leq s \leq 1$) and the elliptic regularity estimate (1.11) remains valid. Remark 1.1 also holds for these problems.

6.2. Stability and Error Estimates. Let $\mathcal{B}_h(\cdot, \cdot)$ be the bilinear form on $V_h \times Q_h$ defined by

$$(6.13) \quad \mathcal{B}_h((\mathbf{v}, q), (\mathbf{w}, r)) = a(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, q) + b(\mathbf{v}, r) - c_h(q, r).$$

Lemma 6.3. *The stability estimate*

$$(6.14) \quad \|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)} \approx \sup_{(\mathbf{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}_h((\mathbf{v}, q), (\mathbf{w}, r))}{\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \quad \forall (\mathbf{v}, q) \in V_h \times Q_h$$

holds for sufficiently small h .

Proof. Let S_h be the supremum on the right-hand side of (6.14). It suffices to show that

$$(6.15) \quad \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim S_h \quad \forall (\mathbf{v}, q) \in V_h \times Q_h,$$

since the opposite estimate follows from the results in Section 2.2 and the Poincaré-Friedrichs inequality [16]

$$(6.16) \quad \|q\|_{L_2(\Omega)} \lesssim \|q\|_{H^1(\Omega; \mathcal{T}_h)} \quad \forall q \in Q_h.$$

For any $(\mathbf{v}, q) \in V_h \times Q_h$ we have, in view of (6.6), (6.13) and (6.16), an obvious estimate

$$(6.17) \quad \begin{aligned} a(\mathbf{v}, \mathbf{v}) &\leq \mathcal{B}_h((\mathbf{v}, q), (\mathbf{v}, -q)) + C_1 \|q\|_{H^1(\Omega; \mathcal{T}_h)} \|q\|_{L_2(\Omega)} \\ &\leq S_h (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) + C_1 \|q\|_{H^1(\Omega; \mathcal{T}_h)} \|q\|_{L_2(\Omega)}. \end{aligned}$$

Let $(\boldsymbol{\zeta}, \theta) \in [L_2(\Omega)]^d \times H_0^1(\Omega)$ satisfy

$$a(\boldsymbol{\zeta}, \mathbf{w}) + b'(\mathbf{w}, \theta) = 0 \quad \forall \mathbf{w} \in [L_2(\Omega)]^d,$$

$$b'(\boldsymbol{\zeta}, r) - c(r, \theta) = \int_{\Omega} qr \, dx \quad \forall r \in H_0^1(\Omega).$$

Then we have, by Remark 6.2,

$$(6.18) \quad \mathcal{B}_h((\mathbf{w}, r), (\boldsymbol{\zeta}, \theta)) = \int_{\Omega} qr \, dx \quad \forall (\mathbf{w}, r) \in [L_2(\Omega)]^d \times H_0^1(\Omega).$$

It follows from the elliptic regularity estimate (1.11) and the interpolation error estimates (2.3), (2.4) and (2.8) that

$$(6.19) \quad \|\boldsymbol{\zeta} - \Pi_h \boldsymbol{\zeta}\|_{L_2(\Omega; \mathcal{T}_h)} + \|\theta - \mathcal{I}_h \theta\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim h^\alpha \|q\|_{L_2(\Omega)},$$

which implies

$$(6.20) \quad \|\Pi_h \boldsymbol{\zeta}\|_{L_2(\Omega)} + \|\mathcal{I}_h \theta\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim \|q\|_{L_2(\Omega)}.$$

We have, by (2.2), (2.19), (6.18)–(6.20),

$$\begin{aligned} \|q\|_{L_2(\Omega)}^2 &= \mathcal{B}_h((\mathbf{v}, q), (\boldsymbol{\zeta}, \theta)) \\ &= \mathcal{B}_h((\mathbf{v}, q), ((\boldsymbol{\zeta} - \Pi_h \boldsymbol{\zeta}), (\theta - \mathcal{I}_h \theta))) + \mathcal{B}_h((\mathbf{v}, q), (\Pi_h \boldsymbol{\zeta}, \mathcal{I}_h \theta)) \\ &\leq C_2 (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) h^\alpha \|q\|_{L_2(\Omega)} + C_3 S_h \|q\|_{L_2(\Omega)}, \end{aligned}$$

and hence

$$(6.21) \quad \|q\|_{L_2(\Omega)} \leq C_2 h^\alpha (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) + C_3 S_h.$$

Combining (6.17) and (6.21), we find

$$(6.22) \quad \|\mathbf{v}\|_{L_2(\Omega)}^2 \leq C_4 S_h (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) + C_5 h^\alpha \|q\|_{H^1(\Omega; \mathcal{T}_h)} (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}).$$

We also have, by (2.2), (2.17) and (6.13),

$$\|q\|_{H^1(\Omega; \mathcal{T}_h)} \lesssim \sup_{\mathbf{w} \in V_h} \frac{b(\mathbf{w}, q)}{\|\mathbf{w}\|_{L_2(\Omega)}} \lesssim \sup_{\mathbf{w} \in V_h} \frac{\mathcal{B}_h((\mathbf{v}, q), (\mathbf{w}, 0))}{\|\mathbf{w}\|_{L_2(\Omega)}} + \|\mathbf{v}\|_{L_2(\Omega)} \lesssim S_h + \|\mathbf{v}\|_{L_2(\Omega)},$$

and hence

$$(6.23) \quad \|q\|_{H^1(\Omega; \mathcal{T}_h)}^2 \leq C_6 (S_h^2 + \|\mathbf{v}\|_{L_2(\Omega)}^2).$$

Putting (6.22) and (6.23) together, we arrive at

$$(6.24) \quad \begin{aligned} \|\mathbf{v}\|_{L_2(\Omega)}^2 + \|q\|_{H^1(\Omega; \mathcal{T}_h)}^2 &\leq C_6 S_h^2 + (1 + C_6) C_4 S_h (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}) \\ &\quad + (1 + C_6) C_5 h^\alpha \|q\|_{H^1(\Omega; \mathcal{T}_h)} (\|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)}). \end{aligned}$$

The estimate (6.15) follows from (6.24) and the inequality of arithmetic and geometric means provided h is sufficiently small. \square

Remark 6.4. The arguments in the proof of Lemma 6.3 are motivated by the arguments of Schatz in [41] for nonsymmetric and indefinite problems.

Similar arguments yield the following stability result.

Lemma 6.5. *The stability estimate*

$$(6.25) \quad \|\mathbf{v}\|_{L_2(\Omega; \mathcal{T}_h)} + \|q\|_{H^1(\Omega; \mathcal{T}_h)} \approx \sup_{(\mathbf{w}, r) \in V_h \times Q_h} \frac{\mathcal{B}_h((\mathbf{w}, r), (\mathbf{v}, q))}{\|\mathbf{w}\|_{L_2(\Omega; \mathcal{T}_h)} + \|r\|_{H^1(\Omega; \mathcal{T}_h)}} \quad \forall (\mathbf{v}, q) \in V_h \times Q_h$$

holds for sufficiently small h .

From now on we assume that (6.14) and (6.25) are valid for all the finite element spaces involved. It follows from these estimates and the same arguments in Section 2.3 that (2.22) and (2.23) also hold for the solution (\mathbf{u}, p) of (6.7)–(6.8) (resp. (6.11)–(6.12)) and the solution (\mathbf{u}_h, p_h) of (6.4)–(6.5) (resp. (6.9)–(6.10)).

6.3. Multigrid Algorithms. The set-up for the multigrid algorithms remains the same, but the definition of the operator $\mathbb{B}_k : V_k \times Q_k \longrightarrow V_k \times Q_k$ is modified as follows:

$$(6.26) \quad [\mathbb{B}_k(\mathbf{v}, q), (\mathbf{w}, r)]_k = \mathcal{B}_k((\mathbf{v}, q), (\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k,$$

where \mathcal{B}_k is the bilinear form on $V_k \times Q_k$ defined by (6.13). The transpose \mathbb{B}_k^t of \mathbb{B}_k with respect to the mesh-dependent inner product $[\cdot, \cdot]_k$ satisfies

$$(6.27) \quad [\mathbb{B}_k^t(\mathbf{v}, q), (\mathbf{w}, r)]_k = \mathcal{B}_k((\mathbf{w}, r), (\mathbf{v}, q)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k.$$

We have the following analog of Lemma 3.3, with an identical proof that uses (6.14) instead of (2.18).

Lemma 6.6. *The norm equivalence*

$$(6.28) \quad [\mathbb{B}_k^t \mathbb{S}_k \mathbb{B}_k(\mathbf{v}, q), (\mathbf{v}, q)]_k^{\frac{1}{2}} \approx \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_k)} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

holds for $k = 0, 1, 2, \dots$

Similar arguments using (6.25) and (6.27) yield another analog of Lemma 3.3.

Lemma 6.7. *The norm equivalence*

$$(6.29) \quad [\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k^t(\mathbf{v}, q), (\mathbf{v}, q)]_k^{\frac{1}{2}} \approx \|\mathbf{v}\|_{L_2(\Omega)} + \|q\|_{H^1(\Omega; \mathcal{T}_k)} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

holds for $k = 0, 1, 2, \dots$

In the definitions of the multigrid algorithms for the problem

$$(6.30) \quad \mathbb{B}_k(\mathbf{v}, q) = (\mathbf{g}, z)$$

arising from (6.4)–(6.5), the pre-smoothing step in (3.14) becomes

$$(6.31) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{S}_k \mathbb{B}_k^t((\mathbf{g}, z) - \mathbb{B}_k(\mathbf{v}_{j-1}, q_{j-1})),$$

and the post-smoothing step in (3.17) becomes

$$(6.32) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{B}_k^t \mathbb{S}_k((\mathbf{g}, z) - \mathbb{B}_k(\mathbf{v}_{j-1}, q_{j-1}))$$

Similarly, in the definitions of the multigrid algorithms for the problem

$$(6.33) \quad \mathbb{B}_k^t(\mathbf{v}, q) = (\mathbf{g}, z)$$

arising from (6.9)–(6.10), the pre-smoothing step in (3.14) becomes

$$(6.34) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{S}_k \mathbb{B}_k ((\mathbf{g}, z) - \mathbb{B}_k^t(\mathbf{v}_{j-1}, q_{j-1})),$$

and the post-smoothing step in (3.17) becomes

$$(6.35) \quad (\mathbf{v}_j, q_j) = (\mathbf{v}_{j-1}, q_{j-1}) + \delta_k \mathbb{B}_k \mathbb{S}_k ((\mathbf{g}, z) - \mathbb{B}_k^t(\mathbf{v}_{j-1}, q_{j-1})).$$

In view of (6.28) and (6.29), we can choose $\delta_k = Ch_k^2$ so that

$$(6.36) \quad \delta_k \cdot \rho(\mathbb{B}_k^t \mathbb{S}_k \mathbb{B}_k) \leq 1 \quad \text{and} \quad \delta_k \cdot \rho(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k^t) \leq 1.$$

6.4. Convergence Analysis. Since the approach is similar we will only point out the necessary modifications and refer to Section 4 and Section 5 for details.

There are now four error propagation operators for the smoothing steps. The error propagation operator for one post-smoothing step is given by

$$(6.37) \quad R_k = Id_k - \delta_k \mathbb{B}_k^t \mathbb{S}_k \mathbb{B}_k,$$

in the case of (6.32), and

$$(6.38) \quad \tilde{R}_k = Id_k - \delta_k \mathbb{B}_k \mathbb{S}_k \mathbb{B}_k^t,$$

in the case of (6.35).

The error propagation operator for one pre-smoothing step is given by

$$(6.39) \quad S_k = Id_k - \delta_k \mathbb{S}_k \mathbb{B}_k^t \mathbb{B}_k,$$

in the case of (6.31), and

$$(6.40) \quad \tilde{S}_k = Id_k - \delta_k \mathbb{S}_k \mathbb{B}_k \mathbb{B}_k^t,$$

in the case of (6.34).

These operators satisfy the following relations:

$$(6.41) \quad \mathcal{B}_k(R_k(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}_k((\mathbf{v}, q), \tilde{S}_k(\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k,$$

$$(6.42) \quad \mathcal{B}_k(S_k(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}_k((\mathbf{v}, q), \tilde{R}_k(\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k.$$

For $0 \leq s \leq 1$, there are two scales of mesh-dependent norms. The norm $\|\cdot\|_{s,k}$ is defined by

$$(6.43) \quad \|(\mathbf{v}, q)\|_{s,k} = [(\mathbb{B}_k^t \mathbb{S}_k \mathbb{B}_k)^s(\mathbf{v}, q), (\mathbf{v}, q)]_k^{\frac{1}{2}} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

and the norm $\|\cdot\|_{s,k}^\sim$ is defined by

$$(6.44) \quad \|(\mathbf{v}, q)\|_{s,k}^\sim = [(\mathbb{B}_k \mathbb{S}_k \mathbb{B}_k^t)^s(\mathbf{v}, q), (\mathbf{v}, q)]_k^{\frac{1}{2}} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

In view of (3.2)–(3.4) and (6.28)–(6.29), the norm equivalences (4.2)–(4.3) also hold for the norms $\|\cdot\|_{s,k}$ and $\|\cdot\|_{s,k}^\sim$ defined by (6.43)–(6.44). Consequently all the results in Section 4.1–Section 4.3 remain valid for these mesh-dependent norms, and in particular,

$$(6.45) \quad \|(\mathbf{v}, q)\|_{s,k} \approx \|(\mathbf{v}, q)\|_{s,k}^\sim \quad \forall (\mathbf{v}, q) \in V_k \times Q_k, \quad k \geq 1.$$

Moreover if we define, for $1 \leq s \leq 2$, the norms $\|\cdot\|_{s,k}$ and $\|\cdot\|_{s,k}^\sim$ by

$$(6.46) \quad \|(\mathbf{v}, q)\|_{s,k} = \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}_k((\mathbf{v}, q), (\mathbf{w}, r))}{\|(\mathbf{w}, r)\|_{2-s,k}} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

$$(6.47) \quad \|(\mathbf{v}, q)\|_{s,k}^\sim = \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}_k((\mathbf{w}, r), (\mathbf{v}, q))}{\|(\mathbf{w}, r)\|_{2-s,k}^\sim} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k,$$

then the results in Section 4.4 also hold for these mesh-dependent norms.

There are now two Ritz projection operators. The operators $P_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ and $\tilde{P}_k^{k-1} : V_k \times Q_k \longrightarrow V_{k-1} \times Q_{k-1}$ are defined by

$$(6.48) \quad \mathcal{B}_k(P_k^{k-1}(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}_k((\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k,$$

$$(6.49) \quad \mathcal{B}_k((\mathbf{v}, q), \tilde{P}_k^{k-1}(\mathbf{w}, r)) = \mathcal{B}_k(I_{k-1}^k(\mathbf{v}, q), (\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k.$$

Property (3.23) remains valid, and it also holds if P_k^{k-1} is replaced by \tilde{P}_k^{k-1} . Consequently we have the following analogs of the Galerkin orthogonality (3.24)

$$\begin{aligned} 0 &= \mathcal{B}_k((Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q), I_{k-1}^k(\mathbf{w}, r)), \\ 0 &= \mathcal{B}_k(I_{k-1}^k(\mathbf{v}, q), (Id_k - I_{k-1}^k \tilde{P}_k^{k-1})(\mathbf{w}, r)), \end{aligned}$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $(\mathbf{w}, r) \in V_{k-1} \times Q_{k-1}$.

Note also that (6.48) and (6.49) imply

$$(6.50) \quad \mathcal{B}_k(I_{k-1}^k P_k^{k-1}(\mathbf{v}, q), (\mathbf{w}, r)) = \mathcal{B}_k((\mathbf{v}, q), I_{k-1}^k \tilde{P}_k^{k-1}(\mathbf{w}, r)) \quad \forall (\mathbf{v}, q), (\mathbf{w}, r) \in V_k \times Q_k.$$

The error propagation operators for the multigrid algorithms are given by (3.22) for the problem (6.30), and

$$(6.51) \quad \tilde{E}_k = \tilde{R}_k^{m_2}(Id_k - I_{k-1}^k \tilde{P}_k^{k-1} + I_{k-1}^k \tilde{E}_{k-1}^p \tilde{P}_k^{k-1}) \tilde{S}_k^{m_1}$$

for the problem (6.33).

Since the proofs of Lemma 5.1 and Lemma 5.2 only involve the results in Section 4 and duality arguments based on elliptic regularity and Galerkin orthogonality, they remain valid for the norms $\|\cdot\|_{s,k}$ and $\|\cdot\|_{s,k}^\sim$ defined in (6.43) and (6.44). Therefore we have the estimates on the effect of post-smoothing coupled with coarse grid correction:

$$(6.52) \quad \|R_k^m(Id_k - I_{k-1}^k P_k^{k-1})(\mathbf{v}, q)\|_{1,k} \lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

$$(6.53) \quad \|\tilde{R}_k^m(Id_k - I_{k-1}^k \tilde{P}_k^{k-1})(\mathbf{v}, q)\|_{1,k}^\sim \lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}^\sim \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.$$

It follows from (4.22), (6.42), (6.45), (6.50) and (6.53) that we have an estimate which measures the effect of pre-smoothing coupled with coarse grid correction for the problem (6.30):

$$\begin{aligned} &\|(Id_k - I_{k-1}^k P_k^{k-1})S_k^m(\mathbf{v}, q)\|_{1,k} \\ &= \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}_k((Id_k - I_{k-1}^k P_k^{k-1})S_k^m(\mathbf{v}, q), (\mathbf{w}, r))}{\|(\mathbf{w}, r)\|_{1,k}} \end{aligned}$$

$$\begin{aligned}
(6.54) \quad &= \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\mathcal{B}_k((\mathbf{v}, q), \tilde{R}_k^m (Id_k - I_{k-1}^k \tilde{P}_k^{k-1})(\mathbf{w}, r))}{\|(\mathbf{w}, r)\|_{1,k}} \\
&\leq \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\|(\mathbf{v}, q)\|_{1,k} \|\tilde{R}_k^m (Id_k - I_{k-1}^k \tilde{P}_k^{k-1})(\mathbf{w}, r)\|_{1,k}}{\|(\mathbf{w}, r)\|_{1,k}} \\
&\approx \sup_{(\mathbf{w}, r) \in V_k \times Q_k} \frac{\|(\mathbf{v}, q)\|_{1,k} \|\tilde{R}_k^m (Id_k - I_{k-1}^k \tilde{P}_k^{k-1})(\mathbf{w}, r)\|_{1,k}^\sim}{\|(\mathbf{w}, r)\|_{1,k}^\sim} \\
&\lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k} \quad \forall (\mathbf{v}, q) \in V_k \times Q_k.
\end{aligned}$$

Similarly the estimate

$$(6.55) \quad \| (Id_k - I_{k-1}^k \tilde{P}_k^{k-1}) \tilde{S}_k^m(\mathbf{v}, q) \|_{1,k}^\sim \lesssim m^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}^\sim \quad \forall (\mathbf{v}, q) \in V_k \times Q_k$$

that measures the effect of pre-smoothing coupled with coarse grid correction for the problem (6.33) follows from (4.22), (6.41), (6.45), (6.50) and (6.52).

Consequently the estimate (5.10) for the two grid algorithm holds for the problem (6.30), and its counterpart

$$\begin{aligned}
(6.56) \quad &\| \tilde{R}_k^{m_2} (Id_k - I_{k-1}^k \tilde{P}_k^{k-1}) \tilde{S}_k^{m_1}(\mathbf{v}, q) \|_{1,k}^\sim \\
&\leq C_* [\max(1, m_1) \max(m_2, 1)]^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}^\sim \quad \forall (\mathbf{v}, q) \in V_k \times Q_k, \quad k \geq 1
\end{aligned}$$

holds for the problem (6.33).

A perturbation argument leads to the following convergence result for the W -cycle algorithm.

Theorem 6.8. *Let E_k (resp. \tilde{E}_k) be the error propagation operator for the k -th level W -cycle algorithm for (6.30) (resp. (6.33)). For any $C_\dagger > C_*$ (the constant in (5.11) and (6.56)), there exists a positive number m_* (independent of k) such that*

$$\begin{aligned}
\|E_k(\mathbf{v}, q)\|_{1,k} &\leq C_\dagger (\max(1, m_1) \max(1, m_2))^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}, \\
\|\tilde{E}_k(\mathbf{v}, q)\|_{1,k}^\sim &\leq C_\dagger (\max(1, m_1) \max(1, m_2))^{-\alpha/2} \|(\mathbf{v}, q)\|_{1,k}^\sim,
\end{aligned}$$

for all $(\mathbf{v}, q) \in V_k \times Q_k$ and $k \geq 1$, provided $\max(1, m_1) \max(1, m_2) \geq m_*$.

7. NUMERICAL RESULTS

We report in this section numerical results that corroborate the theoretical estimates and illustrate the performance of the multigrid methods. The computational domains are the unit square $(0, 1)^2$ and the L -shaped domain $(-1, 1)^2 \setminus [0, 1] \times [-1, 0]$. We use the Raviart-Thomas-Nédélec mixed finite element method of order 1 on uniform meshes in all the numerical experiments, which were supported by the HPC resources of LONI.

7.1. Error in the Nonconforming Energy Norm. In this set of numerical experiments we solve the Darcy system

$$(7.1) \quad \mathbf{u} = -\nabla p \text{ in } \Omega, \quad \nabla \cdot \mathbf{u} = f \text{ in } \Omega, \quad \text{and } p = 0 \text{ on } \partial\Omega,$$

and the convection-diffusion equation

$$(7.2) \quad -\nabla \cdot (\nabla p) + \boldsymbol{\beta} \cdot \nabla p = f \text{ in } \Omega \text{ and } p = 0 \text{ on } \partial\Omega,$$

where $\boldsymbol{\beta} = [2, -1]^T$. We check the error estimate (2.22) by computing

$$\|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)} \quad \text{and} \quad \|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)}.$$

7.1.1. Unit Square. We take the exact solution to be $p = \sin(\pi x) \sin(\pi y)$ and $\mathbf{u} = -\nabla p$. The results are displayed in Table 7.1 and Table 7.2. The index of elliptic regularity $\alpha = 1$ for the square and the convergence rate for $\|p - p_h\|_{H^1(\Omega; \mathcal{T}_h)}$ is 1 for both problems, which agrees with (2.22). The convergence rate for $\|\mathbf{u} - \mathbf{u}_h\|_{L_2(\Omega; \mathcal{T}_h)}$ is 2 for both problems, which is higher than the predicted rate of 1. This is likely due to the phenomenon of superconvergence, since the exact solution is smooth and we use uniform meshes.

TABLE 7.1. Convergence rates for (7.1) on the unit square

h	$\ \mathbf{u} - \mathbf{u}_h\ _{L_2(\Omega; \mathcal{T}_h)}$	rate	$\ p - p_h\ _{H^1(\Omega; \mathcal{T}_h)}$	rate
1/4	4.213e-2		3.462e-1	
1/8	1.026e-2	2.038	1.722e-1	1.008
1/16	2.529e-3	2.020	8.594e-2	1.003
1/32	6.281e-4	2.010	4.297e-2	1.000
1/64	1.565e-4	2.005	2.149e-2	1.000
1/128	3.905e-5	2.002	1.074e-2	1.000

TABLE 7.2. Convergence rates for (7.2) on the unit square

h	$\ \mathbf{u} - \mathbf{u}_h\ _{L_2(\Omega; \mathcal{T}_h)}$	rate	$\ p - p_h\ _{H^1(\Omega; \mathcal{T}_h)}$	rate
1/4	9.188e-2		3.492e-1	
1/8	2.345e-2	1.970	1.727e-1	1.016
1/16	5.904e-3	1.990	8.600e-2	1.006
1/32	1.480e-3	1.996	4.298e-2	1.001
1/64	3.705e-4	1.998	2.150e-2	1.000
1/128	9.268e-5	1.999	1.075e-2	1.000

7.1.2. *L-Shaped Domain.* We take the exact solution to be $p = (1-x^2)(1-y^2)r^{2/3} \sin((2/3)\theta)$ and $\mathbf{u} = -\nabla p$, where (r, θ) are the polar coordinates. The index of elliptic regularity α can be any number $< \frac{2}{3}$ for the L -shaped domain and the exact solution has the correct singularity. The results are presented in Table 7.3 and Table 7.4, which agree with (2.22).

TABLE 7.3. Convergence rates for (7.1) on the L -shaped domain

h	$\ \mathbf{u} - \mathbf{u}_h\ _{L_2(\Omega; \mathcal{T}_h)}$	rate	$\ p - p_h\ _{H^1(\Omega; \mathcal{T}_h)}$	rate
1/4	1.121e-1		1.131e-1	
1/8	6.654e-2	0.752	5.964e-2	0.923
1/16	4.144e-2	0.683	3.244e-2	0.879
1/32	2.605e-2	0.670	1.818e-2	0.835
1/64	1.640e-2	0.667	1.048e-2	0.795
1/128	1.033e-2	0.667	6.186e-3	0.760

TABLE 7.4. Convergence rates for (7.2) on the L -shaped domain

h	$\ \mathbf{u} - \mathbf{u}_h\ _{L_2(\Omega; \mathcal{T}_h)}$	rate	$\ p - p_h\ _{H^1(\Omega; \mathcal{T}_h)}$	rate
1/4	1.306e-1		1.156e-1	
1/8	7.025e-2	0.894	6.023e-2	0.941
1/16	4.231e-2	0.731	3.260e-2	0.886
1/32	2.629e-2	0.687	1.823e-2	0.838
1/64	1.647e-2	0.674	1.049e-2	0.797
1/128	1.035e-2	0.670	6.191e-3	0.761

7.2. **Convergence of Multigrid Methods.** In this set of experiments we carry out the symmetric W -cycle and V -cycle algorithms with m pre-smoothing and m post-smoothing steps for the Darcy system (7.1) and the convection-diffusion equation (7.2). We use the multigrid $V(4, 4)$ algorithm for interior penalty methods to generate the preconditioner L_k in (3.8)–(3.9). We report the contraction numbers obtained by computing the largest eigenvalue of the error propagation operators. The mesh size at level k is 2^{-k} .

7.2.1. *Unit Square.* The contraction numbers of the W -cycle algorithms for (7.1) and (7.2) for various m and $k = 1, \dots, 6$ are presented in Table 7.5 and Table 7.6. For both problems the asymptotic decay rate of $1/m$ for the contraction number predicted by Theorem 5.4 and Theorem 6.8 is observed. (The index of elliptic regularity α for the unit square is 1.) This is also confirmed by the log-log graph in Figure 7.1, where the contraction number of the W -cycle algorithm for (7.1) is plotted against the number of smooth steps m .

TABLE 7.5. Contraction numbers of the W -cycle multigrid method for (7.1) on the unit square

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.80	0.81	0.81	0.81	0.81	0.81
$m = 20$	0.66	0.67	0.67	0.67	0.67	0.67
$m = 40$	0.47	0.48	0.48	0.48	0.48	0.48
$m = 80$	0.24	0.24	0.24	0.24	0.24	0.24

TABLE 7.6. Contraction numbers of the W -cycle multigrid method for (7.2) on the unit square

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.80	0.81	0.81	0.81	0.81	0.81
$m = 20$	0.67	0.68	0.67	0.67	0.67	0.67
$m = 40$	0.48	0.48	0.48	0.48	0.48	0.48
$m = 80$	0.24	0.24	0.24	0.24	0.24	0.25

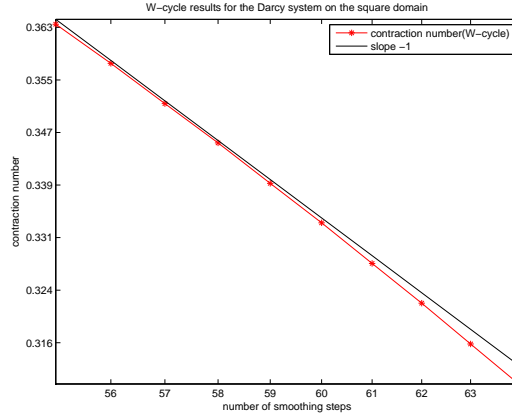


FIGURE 7.1. Contraction numbers of the W -cycle multigrid method for (7.1) on the unit square

We also report the contraction numbers of the V -cycle algorithm for (7.1) and (7.2) in Table 7.7 and Table 7.8. They are similar to the contraction numbers for the W -cycle algorithm in Table 7.5 and Table 7.6 but are slightly larger.

7.2.2. L -Shaped Domain. The contraction numbers of the W -cycle algorithms for (7.1) and (7.2) for various m and $k = 1, \dots, 6$ are displayed in Table 7.9 and Table 7.10. The contraction numbers are larger than the corresponding contraction numbers for the unit square, which is consistent with Theorem 5.4 and Theorem 6.8 since the index of elliptic regularity

TABLE 7.7. Contraction numbers of the V -cycle multigrid method for (7.1) on the unit square

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.80	0.82	0.81	0.82	0.82	0.82
$m = 20$	0.66	0.68	0.68	0.68	0.68	0.68
$m = 40$	0.47	0.48	0.48	0.48	0.48	0.48
$m = 80$	0.24	0.25	0.25	0.25	0.25	0.25

TABLE 7.8. Contraction numbers of the V -cycle multigrid method for (7.2) on the unit square

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.80	0.82	0.82	0.82	0.82	0.82
$m = 20$	0.67	0.68	0.68	0.68	0.68	0.68
$m = 40$	0.48	0.49	0.49	0.49	0.49	0.49
$m = 80$	0.24	0.25	0.25	0.25	0.25	0.25

α for the L -shaped domain is less than $\frac{2}{3}$. This is also confirmed by the log-log graph in Figure 7.2, where the contraction number of the W -cycle algorithm for (7.1) is plotted against the number of smoothing steps m .

The contraction numbers for the V -cycle algorithm are similar and therefore not reported.

TABLE 7.9. Contraction numbers of the W -cycle multigrid method for (7.1) on the L -shaped domain

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.81	0.82	0.82	0.82	0.82	0.82
$m = 20$	0.70	0.70	0.70	0.70	0.70	0.70
$m = 40$	0.51	0.51	0.51	0.51	0.51	0.51
$m = 80$	0.28	0.28	0.28	0.28	0.28	0.28

TABLE 7.10. Contraction numbers of the W -cycle multigrid method for (7.2) on the L -shaped domain

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$m = 10$	0.81	0.82	0.82	0.82	0.82	0.82
$m = 20$	0.70	0.70	0.70	0.70	0.70	0.70
$m = 40$	0.52	0.52	0.52	0.52	0.52	0.52
$m = 80$	0.29	0.29	0.29	0.29	0.29	0.29

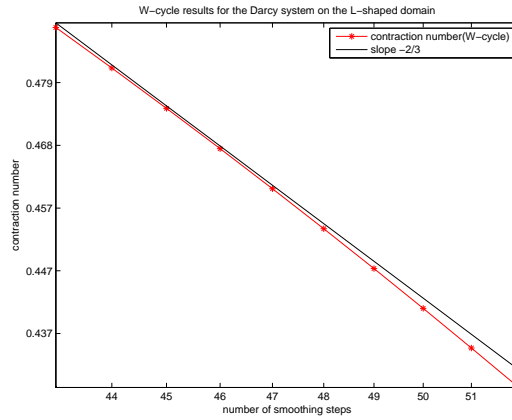


FIGURE 7.2. Contraction numbers of the W -cycle multigrid method for (7.1) on the L -shaped domain

8. CONCLUDING REMARKS

In this paper we developed multigrid algorithms for the Darcy system discretized by Raviart-Thomas-Nédélec mixed finite element methods of order at least 1, and showed that with minimal modifications the multigrid algorithms can also be applied to convection-diffusion-reaction and advection-diffusion-reaction problems. Note that the number of degrees of freedom of the Raviart-Thomas-Nédélec mixed finite element method of order 1 associated with a triangulation \mathcal{T}_h is less than the number of degrees of freedom of the Raviart-Thomas-Nédélec mixed finite element method of order 0 associated with the triangulation $\mathcal{T}_{h/2}$ obtained from \mathcal{T}_h by uniform refinement. Therefore, from the point of view of multigrid, the requirement that the order of the method has to be at least 1 is not restrictive.

The results in this paper can be extended to rectangular Raviart-Thomas-Nédélec mixed finite element methods and to other stable mixed finite element methods for the Darcy system. It should also be possible to extend our approach to mixed finite element methods for linear elasticity that are based on a stress-displacement formulation (cf. [10] and the references therein). We note that a nonstandard analysis similar to the one in Section 5 has been carried out in [42] for a family of such finite element methods.

Finally it would be interesting to extend our approach to the upwind mixed finite element methods for convection dominated problems in [32] (cf. also [33]).

Acknowledgement The authors would like to thank Monique Dauge for helpful discussions on the regularity of the Darcy system.

REFERENCES

- [1] M. B. Allen, R.E. Ewing, and P. Lu. Well-conditioned iterative schemes for mixed finite-element models of porous-media flows. *SIAM J. Sci. Statist. Comput.*, 13:794–814, 1992.

- [2] P.F. Antonietti and B. Ayuso de Dios. Schwarz domain decomposition preconditioners for discontinuous Galerkin approximations of elliptic problems: non-overlapping case. *M2AN Math. Model. Numer. Anal.*, 41:21–54, 2007.
- [3] D.N. Arnold and F. Brezzi. Mixed and nonconforming finite element methods: implementation, post-processing and error estimates. *R.A.I.R.O Modél. Math. Anal. Numér.*, 19:7–32, 1985.
- [4] D.N. Arnold, F. Brezzi, B. Cockburn, and L.D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39:1749–1779, 2001/02.
- [5] I. Babuška, J. Osborn, and J. Pitkäranta. Analysis of mixed methods using mesh dependent norms. *Math. Comp.*, 35:1039–1062, 1980.
- [6] I. Babuška. The finite element method with Lagrange multipliers. *Numer. Math.*, 20:179–192, 1973.
- [7] R.E. Bank and T.F. Dupont. An optimal order process for solving finite element equations. *Math. Comp.*, 36:35–51, 1981.
- [8] F. Ben Belgacem and S.C. Brenner. Some nonstandard finite element estimates with applications to 3D Poisson and Signorini problems. *Electron. Trans. Numer. Anal.*, 12:134–148, 2001.
- [9] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.
- [10] D. Boffi, F. Brezzi, and M. Fortin. *Mixed Finite Element Methods and Applications*. Springer, Heidelberg, 2013.
- [11] J.H. Bramble and S.R. Hilbert. Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation. *SIAM J. Numer. Anal.*, 7:113–124, 1970.
- [12] J.H. Bramble and J. Xu. Some estimates for a weighted L^2 projection. *Math. Comp.*, 56:463–476, 1991.
- [13] J.H. Bramble and X. Zhang. The Analysis of Multigrid Methods. In P.G. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, VII*, pages 173–415. North-Holland, Amsterdam, 2000.
- [14] S.C. Brenner. A multigrid algorithm for the lowest-order Raviart-Thomas mixed triangular finite element method. *SIAM J. Numer. Anal.*, 29:647–678, 1992.
- [15] S.C. Brenner. Convergence of nonconforming multigrid methods without full elliptic regularity. *Math. Comp.*, 68:25–53, 1999.
- [16] S.C. Brenner. Poincaré-Friedrichs inequalities for piecewise H^1 functions. *SIAM J. Numer. Anal.*, 41:306–324, 2003.
- [17] S.C. Brenner, H. Li, and L.-Y. Sung. Multigrid methods for saddle point problems: Stokes and Lamé systems. *Numer. Math.*, published online January 2014. (DOI: 10.1007/s00211-014-0607-3).
- [18] S.C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods (Third Edition)*. Springer-Verlag, New York, 2008.
- [19] S.C. Brenner and J. Zhao. Convergence of multigrid algorithms for interior penalty methods. *Appl. Numer. Anal. Comput. Math.*, 2:3–18, 2005.
- [20] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. *RAIRO Anal. Numér.*, 8:129–151, 1974.
- [21] P.G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.
- [22] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *RAIRO Anal. Numér.*, 7:33–75, 1973.
- [23] M. Dauge. *Elliptic Boundary Value Problems on Corner Domains*, Lecture Notes in Mathematics 1341. Springer-Verlag, Berlin-Heidelberg, 1988.
- [24] J. Douglas, Jr. and J.E. Roberts. Mixed finite element methods for second order elliptic problems. *Mat. Apl. Comput.*, 1:91–103, 1982.
- [25] T. Dupont and R. Scott. Polynomial approximation of functions in Sobolev spaces. *Math. Comp.*, 34:441–463, 1980.
- [26] R. Ewing and J. Wang. Analysis of multilevel decomposition iterative methods for mixed finite element methods. *M²AN Math. Modelling and Numer. Anal.*, 28:377–398, 1994.

- [27] R.S. Falk and J.E. Osborn. Error estimates for mixed methods. *RAIRO Anal. Numér.*, 14:249–277, 1980.
- [28] X. Feng and O.A. Karakashian. Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems. *SIAM J. Numer. Anal.*, 39:1343–1365, 2001.
- [29] J. Gopalakrishnan and G. Kanschat. A multilevel discontinuous Galerkin method. *Numer. Math.*, 95:527–550, 2003.
- [30] P. Grisvard. *Elliptic Problems in Non Smooth Domains*. Pitman, Boston, 1985.
- [31] W. Hackbusch. *Multi-grid Methods and Applications*. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1985.
- [32] J. Jaffré. Décentrage et éléments finis mixtes pour les équations de diffusion-convection. *Calcolo*, 21:171–197, 1984.
- [33] D. Kim and E.-J. Park. A posteriori error estimators for the upstream weighting mixed methods for convection diffusion problems. *Comput. Methods Appl. Mech. Engrg.*, 197:806–820, 2008.
- [34] C. Lasser and A. Toselli. An overlapping domain decomposition preconditioner for a class of discontinuous Galerkin approximations of advection-diffusion problems. *Math. Comp.*, 72:1215–1238, 2003.
- [35] K.-A. Mardal and R. Winther. Preconditioning discretizations of systems of partial differential equations. *Numer. Linear Algebra Appl.*, 18:1–40, 2011.
- [36] S.A. Nazarov and B.A. Plamenevsky. *Elliptic Problems in Domains with Piecewise Smooth Boundaries*. de Gruyter, Berlin-New York, 1994.
- [37] J.-C. Nédélec. Mixed finite elements in \mathbf{R}^3 . *Numer. Math.*, 35:315–341, 1980.
- [38] P.-A. Raviart and J.M. Thomas. A mixed finite element method for second order elliptic problems. In I. Galligani and E. Magenes, editors, *Mathematical Aspects of the Finite Element Method, Lecture Notes in Mathematics 606*, pages 292–315. Springer-Verlag, Berlin, 1977.
- [39] T. Rusten, P.S. Vassilevski, and R. Winther. Interior penalty preconditioners for mixed finite element approximations of elliptic problems. *Math. Comp.*, 65:447–466, 1996.
- [40] T. Rusten and R. Winther. A preconditioned iterative method for saddlepoint problems. *SIAM J. Matrix Anal. Appl.*, 13:887–904, 1992. Iterative methods in numerical linear algebra (Copper Mountain, CO, 1990).
- [41] A. Schatz. An observation concerning Ritz-Galerkin methods with indefinite bilinear forms. *Math. Comp.*, 28:959–962, 1974.
- [42] R. Stenberg. A family of mixed finite elements for the elasticity problem. *Numer. Math.*, 53:513–538, 1988.
- [43] L. Tartar. *An Introduction to Sobolev Spaces and Interpolation Spaces*. Springer, Berlin, 2007.
- [44] P.S. Vassilevski. *Multilevel block factorization preconditioners*. Springer, New York, 2008.

SUSANNE C. BRENNER, DEPARTMENT OF MATHEMATICS AND CENTER FOR COMPUTATION AND TECHNOLOGY, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803

E-mail address: brenner@math.lsu.edu

DUK-SOON OH, DEPARTMENT OF MATHEMATICS AND CENTER FOR COMPUTATION AND TECHNOLOGY, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803

E-mail address: duksoon@cct.lsu.edu

LI-YENG SUNG, DEPARTMENT OF MATHEMATICS AND CENTER FOR COMPUTATION AND TECHNOLOGY, LOUISIANA STATE UNIVERSITY, BATON ROUGE, LA 70803

E-mail address: sung@math.lsu.edu